DOCUMENT RESUME

ED 161 455 IR 006 640

TITLE	The Value of Information: Collection of Papers Presented at 6th Mid-Year Meeting of the American Society for Information Science.
INSTITUTION	American Society for Information Science, Washington, D.C.
PUB DATE	May 77
NOTE	222p.
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Conference Reports; Data Bases; Graphs; *Information Retrieval; *Information Science; Library Automation; *Library Science; On Line Systems; *Professional Associations; Tables (Data)

ABSTRACT

This collection of 39 papers represents the majority of presentations made at the ASIS Mid-Year Meeting at Syracuse University, May 19-21, 1977. Contents are arranged in alphabetical order by first author. The number preceding the author's name coincides with the numbered sequence of presentations in the program bulletin for the meeting. This collection was assembled for the convenience of the participants at the conference to facilitate information interchange. (Author/JPF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161455

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE VALUE OF INFORMATION

****

Collection of Papers

Presented at

6th Mid-Year Meeting

AMERICAN SOCIETY FOR INFORMATION SCIENCE

May 19-21, 1977

Syracuse University

Syracuse, N.Y.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Robert S. Taylor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR006640

INTRODUCTION

This collection of papers represents the majority of presentations given at the ASIS Mid-Year Meeting at Syracuse University, May 19-21, 1977. The contents are arranged in alphabetical order by first author. The number preceding the author's name coincides with the numbered sequence of presentations in the Program bulletin for the meeting.

As valuable as these written documents are, the live presentations and the discussions which followed are the only true "proceedings" of the meeting. This collection was assembled for the convenience of the participants at the conference, to facilitate information interchange. It is not an official document of the American Society for Information Science and should not be considered a "publication" in the usual sense of the word. Copies of this collection will not be available for distribution after the meeting.

We wish to acknowledge with great appreciation the contribution that Tom Morey at Xerox Corporation has made in making this collection possible.

--- Pauline Atherton,
ASIS Mid-Year 1977
Program Chairman

THE VALUE OF INFORMATION -- COLLECTION OF PAPERS

CONTENTS
(arranged by first author)

(Number opposite author's name refers to place in Program bulletin.)

		Page
11 --	Arnovick, George N. The Center for Information and Communication Studies: Criteria and Curriculum for an Innovative Academic Program	1-1J
65 --	Ashton, Carol A., Jean Daubenmire, and Sharon S. Searles An Information Data Base for the Study of Nurse-Patient-Physician Interactions	2-5
61 --	Batts, J.R. and C.F. Powers An Information Delivery System for Survey Research: The Automated Survey Control System (ASCS)	6-11
64 --	Bishop, Ethelyn F. and Audrey Clayton An Application of Market Research Techniques to the Dissemination of Scientific and Technical Information	12-16
43 --	Bohnert, Lea M. A Model to Predict Groups Most Likely to Pay for "Information"	17-20
47 --	Cape, James D. The Value of Energy Information	21-23
49 --	Carrow, Deborah. See Nugent, Joan	
64 --	Clayton, Audrey. See Bishop, Ethelyn F.	
19 --	Clayton, Audrey. See Nisenoff, Norman	
51 --	Collette, A.D. and J. A. Price. A Cost/Benefit Evaluation of Online Interactive Bibliographic Searching in a Research and Engineering Organization	24-34
63 --	Corte, Arthur B. The Importance of EFTS in the Development of Information Systems for the General Public	35-40
20 --	Cox, W.G., Salinshick, F.K. and Wigington, R.L. Cost Effectiveness of Installation of an On-Line Editing System at Chemical Abstracts Service	40A
33 --	Curtis, Dade T. Has the End User Been Forgotten? An Information Profile of Active Research Bioscientists	41-45
65 --	Daubenmire, Jean. See Ashton, Carol A.	

		Page
14 --	Diener, Richard A.V. The Value of Scientific and Technical Information to Developing Areas and Problems in Technology	46-52
68 --	Diener, Richard A.V. Value: Data, Information, Knowledge	53-54
35 --	Dominick, Wayne D. Numeric Data Base Systems: Concepts and Capabilities	55-60
60 --	Featheringham, Tom Present and Potential Value of Computer Communications in Information Science	61
37 --	Freeman, James E. See Kottenstette, James P.	
62 --	Garson, Lorrin R. and Seldon W. Terrant Serving Specialized Subject Interests Among Chemists	62-67
70 --	Glickert, Peter Classification - Human - Value of	68-73
13 --	Hasenjager, Karen. See White, Herbert S.	
32 --	Johnson, F. Douglas The NASA Tech Brief Program: A Cost Benefit Evaluation	74
23 --	King, M.F. See Komoroske, P.A.	
46 --	Kochtanek, Thomas R. Basic Issues in Designing a Pricing Structure for Library and Information Services	75-79
23 --	Komoroske, P.A. and M.F. King Costs of On-Line Bibliographic Searches in Industry	80-86
37 --	Kottenstette, James P. and James E. Freeman Accessing Numeric Data Via Flags and Tags: Progress Report on a Real-World Experiment	87-90
69 --	Levine, Marilyn M. Benefits of Information	91-92
56 --	Linehan, Ronald Statewide Planning for Library Development -- The Information Gap	93-96
53 --	Long, Susan SMERC: An On-Demand Educational Information Center	97-102
50 --	Lundeen, Gerald Introducing the Concepts of On-Line Boolean Searching with a Mini-Computer Based System	103-106

		Page
29 --	McGrath, William E. Efficiency of Competing Models for Predicting Book Circulation by Subject: A Theoretical and Empirical Basis for Allocation	107-113
59 --	Meincke, Peter P.M. World Trade in Information and Know-How	114-119
42 --	Miller, Betty and David Mindel A Corporate Library's Experience with Selling On-Line Information Retrieval	120-125
42 --	Mindel, David. See Miller, Betty	
31 --	Musselman, K. and J. Talavage Periodic Management Review of Information Services	126
19 --	Nisenoff, Norman and Audrey Clayton Information Costs and Values -- The Next 15 Years	165
49 --	Nugent, Joan and Deborah Carrow Ins and Outs of an In-House Evaluation - Problems and Solutions from a Working Information System	127-135
28 --	Palmer, Richard Integrating Library Acquisitions with the Curriculum	136-139
61 --	Powers, C.F. See Batts, J.R.	
51 --	Price, J.A. See Collette, A.D.	
24 --	Rathswohl, Eugene J. Educating Management Students in Information Utilization	140-146
48 --	Roberts, Suzanne H. Information and Data System Simulation A Report of 1976 Experiences	148-155
25 --	Rubin, Stephen A. Service Charging: An IAC Experience	156-163
65 --	Searles, Sharon S. See Ashton, Carol A.	
21 --	Sherman, Philip M. Costs and Values of Technical Information in Industry	164
66 --	Snodgrass, Rex J. See Wilde, Daniel U.	
18 --	Somerville, Arleen N. Costs and Value of Scientific and Technical Information in a University Environment	166-167
26 --	Stewart, Alan K. Equipment Selection Criteria for 1200 Baud Bibliographic Searching	168-173

Page

62 -- Terrant, Seldon W. See Garson, Lorrin R.

13 -- White, Herbert S. and Karen Hasenjager
The Library Doctorate - A Self-Assessment of Degree Holders 174-194

66 -- Wilde, Daniel U. and Rex Snodgrass
The Value of Searching Multiple Data Bases 195-199

THE CENTER FOR INFORMATION AND COMMUNICATION STUDIES: CRITERIA AND CURRICULUM FOR AN INNOVATIVE ACADEMIC PROGRAM

George N. Arnovick
Professor, Department of Computer Science
California State University, Chico
Chico, CA 95929

INTRODUCTION

The Center for Information and Communication Studies (CICS) was implemented in the Fall of 1974, and was the result of an extensive study undertaken by a special Task Force. The members of the study group were faculty from mass communications, business, psychology, education, computer science, library, and instructional media. The disciplines represented by the above faculty were considered necessary, as the concept of interdisciplinary involvement required an integration of different viewpoints and academic considerations that are germain to the broad functions within information and communications.

The planning, study, and outgrowth of a Center concerned with teaching and research in the information and communications area was a logical outgrowth of programmatic moves made by the University in 1973 which lead to the establishment of the Learning Activities Resource Center (LARC). LARC is a service unit of the campus which consolidated print and non-print materials and mediated systems. The bringing together of the library and instructional media was a deliberate step to integrate the process of combined forms of knowledge and distribution so that students, faculty, and administrative functions could be aided by having a consolidated resource. During the 1976-77 academic year and at the present time, LARC includes the University Computer Center which previously reported to the office of the Vice President for Administration. Since the Computer Center services academic and administrative needs, the locating of computer operations within a learning resource environment facilitates the information retrieval requirements of the University.

CSUC, as one of the nineteen campuses within the California State University and College system, has a unique distinction inasmuch as the campus is located in North East California, and has a service area approximately the size of the state of Ohio. The service area represents foothills and mountain communities which are widely separated and very difficult to reach during the winter season due to snow conditions. As part of the Northern California Regional Programs Consortium, which includes community colleges, CSUC subscribes to the concept of the "Open University" in terms of delivering education. LARC, with its mediated delivery system and micro-wave, including computer services, has the initial capability of enabling education and information to reach students in an extended campus environment. CSUC is the first campus within the system to have a Dean of Learning Resources. This background is provided regarding LARC so that an understanding and persceptive is gained on how the new CICS is the academic arm of LARC.

GOALS AND OBJECTIVES

The primary objective for the design of curriculum and the structure of the Center was to meet the need and challenge of an emergent information/communications society. New careers and educational criteria brought about by technological and social change indicated that academic goals have to be developed to allow students to study and become proficient in information and communication processes to meet career opportunities.

To meet the educational goals which could provide academic emphasis on understanding and interpreting the new convergent disciplines of information and communications required significant change and upgrading of courses available to students at the University. The department of mass communications was identified as being the most appropriate academic area that by discipline and course offerings touched on the requirements for a new curriculum. Since the department of mass communications had a liberal arts & technology emphasis, it was decided that the new center would be most effective by reconstituting the mass communications program.

As determined by the study, the following objectives for the Center were chosen for the new program:

1. To develop an implementation plan and recommendations for the initiation and operation of an interdisciplinary program unit concerned with teaching and research in the area of Information and Communication Studies for California State University, Chico.

2. To acknowledge present and new educational settings and subject matter so that a viable and creative union results among divergent academic areas combining the humanistic, artistic, and scientific aspects of those identified academic disciplines as they apply to the Information-Communication program.

3. To provide an educational setting for faculty-student interaction which results in;

   - Core studies broadly covering the structure, processes, functions, impacts, and theories of media communication, information science, and instructional technology.

   - Increased skills centered around production courses interconnecting various software and hardware resources and systems.

   - Services to the University service area which encourage and support the effective utilization of media systems and procedures in the resolution of societal problems and issues and in revitalization of our institutions.

   - Research skills which utilize communication activities and resources to continually investigate messages, channels and activities, characteristics of media agencies and institutions regionally and at the state and national levels.

° Provision for the development of understandings, competencies, and social awareness, in the utilization and teaching the use of all print media, nonprint media and equipment as they relate to educational goals, the curriculum and individual learning situations.

THE CICS PROGRAM AND CRITERIA

Given the learning goals and objectives as defined previously, an input-output concept was developed. Students were inputs; outputs were the knowledge areas and skills obtained relative to broad social needs and specific career needs. The subject-course area could then be designed as the process which took into consideration student characteristics, study goals, ways of teaching, and units of instruction by identified interdisciplinary subject areas. The process of learning can then be optimized to meet output--those social, environmental, economic and individual factors which make up the relative composite requirements for a program based on the interrelationships of information and communication as an academic concentration.

Characteristics

The program offered by the Center is designed so that students can achieve an education in modern information and communications. The curriculum is planned so that it serves those students enrolled as majors or minors, and for all other students interested in the area of information and communication to satisfy electives and the rounding out of their general academic experience.

The Center can be considered to be the University's focal point in academic responsibility for stimulating the critical appreciation for the nature and significance of modern information processes, media systems, and communications functions.

Surrounding this core responsibility is organized a broad spectrum of related academic disciplines that contribute to the process of academic inquiry aimed at understanding the eclectic fields of information and communications. The nature, functions and effects of modern information handling and communications draws on theory and processes derived from human understanding, technological advances, social and economic change, and the aculturation of society in general.

Technology and man, when coupled in an interactive process, has a common goal--the modification of human conduct. These apply equally to organizational content and to individual behavior. The technology and social impact imposed upon human conduct by the systems and processes of modern information transfer and communications, have reached sophistication of such a degree as to acquire the closest scrutiny.

Communications today, and even more so expected in the future, carry an information channel whose messages reach an ever-growing population. The distinction of objective news and propaganda becomes blurred. What is announced today can be, and has been, exploited to the detriment of society.

With the advent of more automated systems in the future, rights of privacy

from information generated by computers and communicated to multi-source outlets in every conceivable form of transmission will likely be violated to a greater degree than anticipated by current legislation dealing with privacy and rights of data. No longer can such transgression be known only to a few--those who are intimately concerned with information and communications process.

Our society has taken communications for granted. What education must do now is develop conceptual insight and perspective, in addition to skills, so that learning can be achieved by identifying the problem and posing solutions with regard to information transfer and mediated systems. Such knowledge can result in understanding, measuring, and ultimately contributing to human purposes the nature, quality and impact of such systems. In so doing, the Center defines its programs and responsibility in the tradition of liberal education, augmented by those elements which deal with the scientific principles of man-machine systems.

These learning goals involve the fundamental interactions of information science, instructional technology, and public communications. These areas are identified by the center faculty as the basic educational platform on which the CICS curriculum is based.

CURRICULUM

The curriculum rationale is based on deriving an academic program which integrates those broad aspects of information, communications, and instructional technology disciplines. The basic idea is to bring these disciplines into focus as an integrated study approach rather than emphasizing each discipline category separately.

Justification and purpose of the curriculum takes into account the emphasis and importance of the above discipline areas which have been recognized as follows:

> A 1970 Rand Corporation report on communication technology has stated that "the revolution in communication...is just beginning." Where and how communication technology "will eventually fit into more comprehensive communication systems will not simply be a function of technical capabilities and cost considerations but also of public policy and of entrepreneurial and consumer pressures and choices." The report suggests that before the full usefullness of the new developments can be realized, social effects of the new technology, and public policies concerning the utilization of the new technology must be studied.
>
> ----"The Social Effects of Communication Technology", RAND Corporation, May, 1970.

The Carnegie Commission on Higher Education recommended in 1972:

> "Institutions of higher education at all levels should develop their potential for training specialists and professionals needed to perform the new functions that are associated with the increasing utilization of instructional technology on the nation's college and university campuses."

----"The Fourth Revolution, Instructional Technology in Higher Education, Carnegie Commission, June, 1972.

Information science is an emerging field which has developed in response to a national need. "The problem" as McGeorge Bundy stated in a keynote address to a Congressional Committee, "is that in most, if not all spheres of inquiry and choice, quantities of raw information overwhelm in magnitude the few comprehensive and trusted bodies or systems of knowledge that have been perceived and elaborated by man." New systems are required, new methods of handling information, new techniques of managing knowledge. Schools of information science are necessary to educate a new professional class of information analysers, organizers and managers to solve this problem.

----"U.S. Congress, House. Committee on Science and Astronautics. The Management of Information and Knowledge; a Compilation of Papers Prepared for the Eleventh Meeting of the Panel on Science and Technology (Washington, D.C.: Government Printing Office, 1970).

Curriculum Concept

The concept for curriculum offerings within the Center is based on the recommendations as previously discussed in the objectives. The curriculum planning model follows those objectives to create an interdisciplinary information and communication studies program. The basic goal is to provide for a broad and in-depth education, which is based on a mix of liberal arts and sciences, to reflect the different types of course contents, general education, supportive electives, and required co- and prerequisites.

The recommended curriculum structure follows a three-option approach, each being an identified academic area of emphasis. Each of these components is a structured logical unit within the overall course of study concerned with the information and communication studies program. To allow for academic flexibility, a student may choose a special option which is a combination of options within the program. These options are as follows:

- Public Communications
- Information Studies
- Instructional Technology
- Special Option (combination of above)

For choice of concentration, each option has several advising sequences, which are specifically identified areas that the student can emphasize within the broad structure of the option of their choice. An example of such are listed below:

Option	Advising Sequence
1. Public Communication	1. Broadcasting 2. Journalism 3. Visual Communications

2. Information Studies
   1. Organizational
   2. Systems

3. Instructional Technology
   1. Professional Block (examples of emphasis):
      - Industrial/Business Media
      - Computer Systems of Instruction
      - Educational, Administration, Curriculum Development

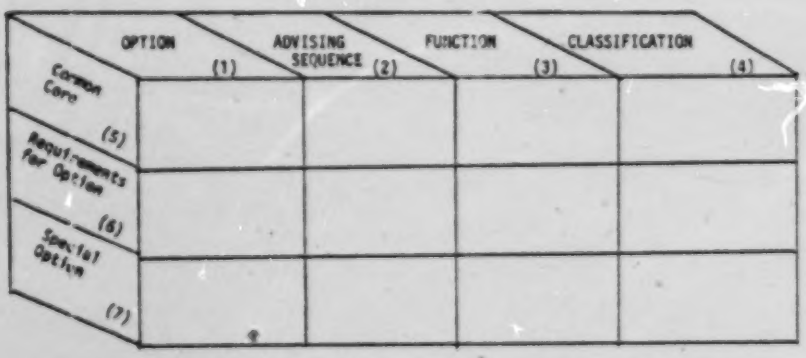
All options follow a classification which is the Center's "system" approach to learning via structured courses. The classifications developed from the overall curriculum are as follows:

- Theoretical--Concepts
- Application--Services
- Research--Technique
- Methods--Forms
- System--Integration

So that courses can be specifically related to topical categories, a function dimension is provided so that related courses can be grouped. These groupings are helpful in modular course offerings as well as identifying curriculum balance within a pattern structure.

The Center's educational approach emphasizes a horizontal (across disciplines) as well as vertical (within disciplines) pattern structure. To accomodate this flexible form of course exposure, all majors within the Center will be required to take courses in each of the three option areas regardless of their pattern of choice. These types of courses, or modular blocks, are identified as the Common Core. Those subjects and associated courses which make up the in-depth learning process within each option are identified as option requirements. Figure 1 shows a three-dimensional matrix of the Center's curriculum concept.

FIGURE 1
CICS Curriculum Dimension



(1) Information Studies, Instructional Technology, Public Communications (discipline area)
(2) Area of concentration within option (specialization)
(3) Theoretical-concept, methods-forms, application-service, system-integration, research-techniques
(4) Topical discipline groups within option (grouping of related courses)
(5) Taken in each of the options
(6) Within the option discipline
(7) Combined option of choice

Curriculum Contents

A summary of the contents of the curriculum is presented in Figure 2. A detailed list of all of the catalog descriptions of courses is not given due to the number of courses which would take up several pages. Those persons interested in obtaining a course catalog can write to the author who will be pleased to send a copy.

FIGURE 2
Summary of Course Areas of Emphasis Within Options

Public Communication Option

Mass Communications:
Radio, TV, Print Media
Visual Communications:
Graphics, Kinegraphics, Visual Communication & Environment, Design, Packaging
Organizational Communication:
Public Information, Advertising, Communication Management
International & Intercultural Communication:
Communications among Nations, Cultures, and Peoples
Political Communication:
Impacts of Communication on Political Behavior; Dissemination of Political Information; Ethics of Conducting Political Campaigns
Interpersonal Communication:
Person-to-Person Communication; Small Group Communication; Communication & Marital Relationship; Nonverbal Communication

Instructional Technology Option

Research & Theory in Instructional Technology
Technology & Modern Education
Use of Instructional Media in Public Schools (& Colleges)
Design and Production of Instructional Materials
Design and Operation of Instructional Systems
Computer Assisted Instructions
Planning and Preparing Instructional Objectives
Learning, Perception, and Mass Communication Theories Applied to Mediated Instruction
Evaluation of Instructional Methods and Materials
Administration of an Instructional Materials Center

Information Studies Option

Informatology:
Interrelationship of disciplines (Math, Psychology, Media, Communication, etc.)
Cybernetics:
(Feedback Control, Man-Machine Communication)
Languages for the Library and Information Sciences
Organizational and Analytical Aspects of Information Sciences:
(systems, information management, concepts)
Technological System:
(design, information representation, modeling, information retrieval, structure and classification)
Interactive Systems:
(Computer courseware, natural languages and human communication, synthetic forms)
Applications and Methods
(Document analysis, computational linguistics, semiotics, pattern recognition, measures of effectiveness, simulation)

The Center has added numerous new courses to reflect the curriculum, as well as modified a number of courses previously taught under mass communications. At the present time, three to four new courses still require University approval via department and school recommendation processes. The Center expects these courses to be approved and offered beginning with the Fall 1977 semester.

DEGREE AND COURSE STATUS

The CICS curriculum and new degree offering has been approved by the Chancellor's Office effective for the Fall 1977 semester. The degree includes all three options. From Fall 1974 when the Center was established until the present (Spring 1977 semester), the B.A. degree and M.A. degree has been granted but the course offerings were limited to the Public Communications Options in the undergraduate program. The M.A. areas of study includes Public Communications and Instructional Technology. With the current B.A. degree approval, the M.A. program will be able to offer courses in Information Studies so that graduate students can concentrate in this area.

Undergraduate Degree

The requirements for the B.A. degree are shown as follows:

Requirements for the B.A. in Information and Communication Studies (CICS): 124 units total

General Studies: 40 units. Students majoring in Information and Communication Studies may not use the Center's courses to satisfy general studies requirements.

Common Core: Requirements for all CICS majors: 18 units

CICS 1A	Principles of Information and Communication	3
CICS 101	Modern Communication Systems and Society	3
CICS 201	Theories of Information and Communication	3
CICS 231	Mass Communication and the Law	3
CICS 255	Informatology	3
CICS 202	Communications Research	3

Option Requirements: In addition to the common core classes, students must also satisfy requirements for one of the following options and its respective advising sequence:

Option	Advising Sequence	Units
Public Communication	Broadcasting	25
	Journalism	24
	Visual Communication	23
Information Studies	Organizational	21
	Systems	21
Instructional Technology	Professional Block	21

Minor or Cluster Requirements: 18-21 units. All CICS majors will develop a formal minor in a department outside of CICS or develop an adviser-approved cluster of courses totaling 18-21 units in appropriate outside fields. For students who elect to choose the Information Studies Option, a minor in Computer Science is required. Certain courses within CICS can be applied to the approved cluster provided such courses are not offered within the student's degree option.

Electives: 20-28 units. Electives to be taken from University offerings.

Figure 3 below is a summary of the number of units required for each option and advising pattern:

FIGURE 3
Summary Unit Table for All Options and
Advising Sequences (B.A. CICS)

OPTIONS	PUBLIC COMMUNICATIONS			INFORMATION STUDIES		INSTRUCTIONAL TECHNOLOGY
ADVISING SEQUENCE	Broadcasting	Journalism	Visual Comm	Organizational	Systems	Professional Block
COMMON CORE	18	18	18	18	18	18
OPTION REQUIREMENTS	13	18	11	6	6	12
REQUIRED ELECTIVES	11-12	6	12	15	15	9
SUB TOTAL	42-43	42	41	39	39	39
CLUSTER OR MINOR REQUIREMENTS OUTSIDE CICS	18-21	18-21	18-21	18-21	18-21	18-21
TOTAL	60-64	60-63	59-62	57-60	57-60	57-60

Graduate Degree

The M.A. program in the Center is broad and interdisciplinary in nature, allowing students a great deal of flexibility in course offerings that deal with information and communications. Specific areas of concentration are mass communications, political communications, intercultural and international communications, organizational or management communication, instructional technology, (information studies commencing Fall 1977), or a combined emphasis of any of the above areas of concentration. Since the degree is broad in content, all graduate candidates are required to take 7 units consisting of seminars in communication studies, quantitative methods in information and communications, and a forum in communication and instructional technology. The overall program requires 30 units of 200/300 level courses (200 = senior/grad, 300 = grad), half of all units must be at the 300 level, and no more than a total of 6 units of independent study, no more than 6 units in each of the non-major areas.

ORGANIZATION

The Center reports to the Dean of Learning Activities Resource Center (LARC) as recommended by the Center's faculty. The faculty decided that since the concept of the Center as an academic unit was non-traditional, reporting to an existing school would not provide the intellectual and support environment that would be in the interests of the program. Since LARC has the media facilities and print resources, as well as the computer center, these resources, when tied into the Center's curriculum, represents a unique opportunity for achieving the

the Center's programmatic objectives. The Center, as part of LARC, represents the academic side of learning resources and compliments the University's outreach program, as well as allowing the Center to provide teaching and research input so that LARC can optimally meet its goals of providing learning resources (equipment, methods, services) to the University.

In terms of traditional faculty relationships and functions pertaining to RTP and academic senate representation, some problems have arisen. However, these have been minor and reflect the necessary changes forthcoming when an academic unit is tied into a service component of the University. An evaluation forthcoming within the next two years will allow the Center to assess itself, and could conceivably be affiliated with a school at such later time if the curriculum and other factors demand a traditional school affiliation. In summary, the organization and faculty participation is an experiment, and as such, needs time to work out solutions which afford the best interests of the program, students, and faculty.

FUTURE PLANS

In terms of curriculum, the Center has already defined major objectives to alternative methods of instruction. This phase will commence in the Fall 1977 semester and will consist of developing thematic and block courses instead of the traditional 1-3 unit structure. Computer aided instruction, individually paced learning modules, and professor-student facilitator roles will be examined and hopefully entered into an experimental basis.

The Center's long range master plans consider a number of innovative approaches to education, which depart from traditional concepts. Since the curriculum and degree programs consider information and communications process, it is strongly anticipated that our program will not only serve our students, but can become a model for other instructional programs in the University.

Considering our FTE growth (almost 10% per year from 1974) and the fact that the Center now has approximately 500 majors (from 267 in Fall 1974), including an FTE rate increase of 4 1/2 times faster than the University's average annual growth, alternative instructional approaches become mandatory if we are to offer an education to our students. Considering that University enrollments are generally declining, and budget cutbacks a way of life, the Center represents an exception in terms of student growth. However, our budget support to include faculty hiring remains a problem. Budgets are generated on an average of an 18:1 student faculty ratio. The Center has an average of 23.5:1. Due to student faculty ratio averaging on a University-wide basis, which takes into account declining enrollments in other areas, the Center cannot grow with the same ratios. We have already declared an impact in the Broadcast Option, and will probably declare student enrollment limits in other options shortly.

On the optimistic side, the Center may be the forerunner of adopting alternative methods of instruction so that unified resources (faculty/budget) can be accomodated without sacrificing educational integrity.

REFERENCES

1. Debbons, A., and others, 1970. "Toward a Metascience of Information: Informatology." Journal of the American Society for Information Science, 1970.

2. Ashby, Eric. 1972. "Fourth Revolution, Instructional Technology in Higher Education," Carnegie Commission on Higher Education, June 1972.

3. Arnovick, G.N. 1974. "Task Force Report on the Center for Information & Communication Studies," California State University, Chico, 1974.

4. Arnovick, G.N., and Pennington, F. 1972. "Final Report of the Task Force on Instructional Media," California State University, Chico, 1972.

AN INFORMATION DATA BASE FOR THE STUDY OF NURSE-PATIENT-PHYSICIAN INTERACTIONS

Carol A. Ashton, R.N., M.S., M. Jean Daubenmire, R.N., M.S., Sharon S. Searles, R.N., M.S., School of Nursing, The Ohio State University, Columbus, Ohio

INTRODUCTION

The communication process is complex and basic to all human relationships. Learning to communicate effectively is important for all persons but becomes even more important for nurses, patients and physicians in a hospital environment. The patient is concerned about his diagnosis and sensitive to both the verbal and the nonverbal messages which occur during an interaction. Very little research has been done on this communication and its effects on nurses, patients and physicians.

Because the interaction process between persons is complex and difficult to study many investigators have attempted to control behavior under experimental conditions only to discover that the findings are not applicable to a "real world" situation. Thus, the purpose of Daubenmire's study is to develop a methodology for examining the complexity of the interaction processes that occur among nurses, patients and physicians in a natural or hospital environment.

BACKGROUND

After a thorough review of the literature, it was apparent that most studies related to communication either dealt with: 1) a single or limited number of variables, 2) the frequency or occurrence of those variables, and 3) controlled or experimental environments. Daubenmire then, has developed an unique methodology which provides a model, a notation system, and an analytical framework to study nurse-patient-physician communication. Although the methodology is applicable to study any nurse-health care personnel communication, the present focus of study is nurse-patient communication.

Communicative interaction is conceptualized as a dynamic, ongoing behavior process which can be studied by identifying the observed effects of presenting behaviors on the participants. Presenting behaviors are those verbal and nonverbal behaviors by which persons communicate with each other. Presenting behaviors are extremely important because a human relationship cannot be directly observed; only the behaviors of individuals can be observed and measured.

One unique aspect of the Daubenmire study is an extensive notation system which is used to describe the quantitative and qualitative aspects of behavior. The quantitative aspect of behavior refers to the occurrence or frequency of behavior. The qualitative aspect refers to the way in which or <u>how</u> a behavior occurs.

The major framework for qualitative description of behavior is adapted from a system of dance notation, effort-shape. So, for example, a touch was coded as occurring and that touch was further described as whether it was a forceful, direct touch or a light and sustained touch.

The notation system then preserves the nurse-patient interactional data in a form and degree of complexity not previously available to investigators.

INFORMATION DATA BASE

A patient room in a Columbus, Ohio hospital was equipped with a remote-controlled, audio-video recording system and time series data of patient-health care personnel interactions were collected throughout a hospitalization from admission to discharge. Four patients were involved in the study and all patients and personnel granted permission for videotaping prior to the actual recording sessions. There are 122 hours of videotape data and 1,902 interactions between patients and different personnel.

Data were extracted from videotapes and, utilizing the afore-mentioned notation system, coded into mnemonic word form and computerized. After this process, the data base consisted of time series data, i.e. a second-by-second record of all occurring verbal and nonverbal behaviors of both nurse and patient throughout an interaction. This data base was then converted to numeric files. Present and future investigators have available mnemonic or readable data as well as a numeric file which greatly facilitates data analysis.

DATA ANALYSIS

A variety of analytical techniques are currently being used to: 1) identify key verbal and nonverbal behaviors, 2) identify patterns of behavior between nurse and patient, and 3) identify initial measures of or rhythmns of convergence and synchrony.

Due to the exploratory and descriptive nature of Daubenmire's research, a grounded theory approach to analysis is being utilized. The approach is classified into three levels: 1) the empirical, 2) the statistical, and 3) the conceptual. The first, and most basic level is the empirical level. At this level, techniques are applied to the extensive data base in an attempt to identify key behaviors or patterns and significant relationships among the behaviors which seem most explanatory or predictive of the interaction.

The statistical level of analysis involves more vigorous testing of the identified key behaviors and relationships among behaviors. The findings from the first two levels of analysis can then be used to further refine, revise, and develop the conceptual framework for study. Concepts and hypotheses can be formulated and cycled into the empirical or statistical levels for further investigation. Therefore, this approach to data analysis is cyclical in nature i.e., findings from each level of analysis provide a basis for interpretation and further development and refinement of the framework.

VALUE OF INFORMATION

The values of such an information data base are many and far ranging. The interactional data is preserved in three forms: 1) videotape, 2) computerized mnemonic word form, and 3) computerized numeric files. Because the data are preserved, many investigators could access these data for replication studies, spin-off studies, and new research studies into specific areas of communication. The data are also available in both observable and readable forms so that an investigator could develop different designs. The data are time series data thus various stochastic or process-oriented analyses could be applied.

The Daubenmire information base is crucial to the study of communication as a process. The computerized data are being analyzed to identify key verbal and nonverbal behaviors and to identify patterns of behavior between nurse and patient. Insights into key behaviors would be of great help to not only health care personnel but also to any individual to whom communication with a client, personnel, a corporation, etc. is critical. In other words, new insights into communication would be of great help to all individuals. For example, findings from this study may validate the difficulties with computer conferencing, i.e. verbal interaction is only one small part of communication; nonverbal interaction and human response are as necessary if not more so for meaningful human interaction.

Identification of key behaviors and prediction of interactional patterns would lead to the development of educational strategies for the health care professions, businesses, management, etc. For example, if a nurse could predict a certain patient behavior pattern which tends to be problematic, the nurse could then identify an intervention strategy to facilitate a more positive change in the patient's pattern. Warren Lamb, who conducts movement studies in England, currently uses movement profiles to predict and select those persons who are most likely to perform well in certain management positions. The future potential exists to match certain teachers with a certain group of students or a certain nurse to a certain group of patients.

From a sponsor cost-benefit perspective, a pre-existent data base such as Daubenmire's would save thousands of sponsor-dollars and that money usually allocated for data collection could be redistributed. The information data base could be used in innumerable ways and by a variety of investigators.

SUMMARY

Although the focus of Daubenmire's study is to examine nurse-patient interaction, the methodology developed and the information data base can be used to study human interaction in any setting. The richness of multivariate data has been preserved and the identification of key concepts in human interaction will provide knowledge useful in all areas of education, practice and research.

REFERENCES

Daubenmire, M. Jean and King, Imogene. "Nursing Process Models: A Systems Approach." Nursing Outlook. Vol. 21, No.8, August, 1973.

Daubenmire, M. Jean. "Caring is Not Enough." Physicians World. Vol. II, No.10, October, 1974.

_____. "Nurse-Patient-Physician Communicative Interaction Process." The Systems Approach in Health Research. New York: Springer Publishing Co., 1976.

Hawes, Leonard C. "Elements of a Model for Communication Processes." The Quarterly Journal of Speech. Vol. 59, 1973.

AN INFORMATION DELIVERY SYSTEM FOR SURVEY RESEARCH: THE AUTOMATED SURVEY CONTROL SYSTEM (ASCS)

J.R. Batts and C.F. Powers
Research Triangle Institute
Research Triangle Park, North Carolina 27709

INTRODUCTION

Traditionally, survey management and control have been performed by project staff maintaining, on a day to day basis, some form of manual information delivery/reporting system. Such systems require written documentation and reporting of survey events and can be supplemented with visual aids which provide an overview or summary of event status. When applied to large, complex surveys, the manual system often requires the maintenance of multiple accounting documents that tend to quickly become fragmented operationally from the survey. The daily monitoring of survey activities becomes cumbersome because of the cross-referencing required between the various accounting documents. Inefficiency in allocating project tasks results because the system cannot provide an accurate update of survey events. The manual control system itself eventually becomes dysfunctional to the survey, requiring more time and attention than the survey's operational aspects.

It is because of the limitations of the manual control system that an alternative to survey management was developed. This paper reports the design and development of an automated survey information delivery system (1) that was developed to streamline the overall management process and thereby reduce the management burden of survey research without reducing operational control. The Automated Survey Control System (ASCS) evolved from a need to bring together integral parts of survey support under one basic efficient and accurate control mechanism - the computer. As is oftentimes the case with large scale surveys, it is imperative that the status of many field operations be monitored with accuracy and a quick reporting capability. Commonly used management techniques reveal their inadequacies where such needs are realized. Thus it was the need for accuracy and reporting that prompted the bringing together of many elements of the survey control effect to formulate a new system.

OPERATIONAL OVERVIEW

The concept of computerized data management systems is certainly not new (2). Indeed, many of the components of the system described in this paper have been utilized separately on many survey research projects in both manual and automated form. However, the ASCS is the result of the combination of existing manual and automated techniques with new ideas to create an automated survey support system.

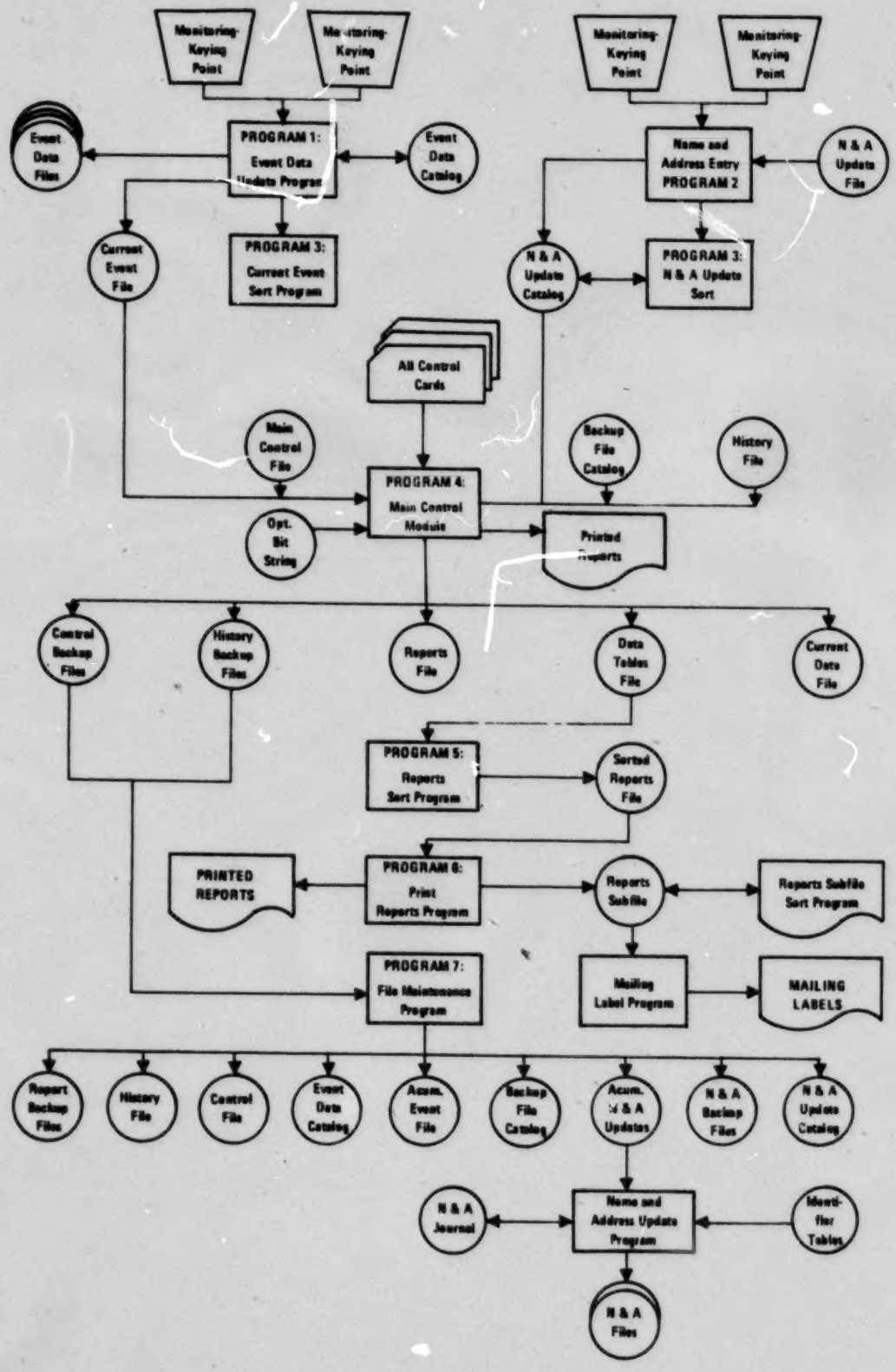


Figure 1. ASCS Flow Chart

Generally, the ASCS provides project management with reports of activities reflecting the current status of the survey at large and/or any subset of the survey. Daily monitoring of survey activities as reported by project staff through received data are reflected in the reports as well as cross tabulations of each of the operational categories (events). The flow chart in Figure 1 gives an overview of the data flow of a typical survey operation. There are several important elements of data flow that appear in the flow chart that are not discussed in this paper. These are the quality control software, machine editing software, and the error correction steps. The flow chart shows how each of these project elements coordinate with the ASCS. Basically, the ASCS accepts externally generated event data, an endogenous data table describing the survey event structure, and report requests and uses this information to update the control file. The control file contains a record for each sample participant in the survey and each record is accessed on each run, even though in many cases no external input affects the given record. A record may, however, be modified based upon information contained in the data table describing the events.

The primary ASCS program (Program 4 of the flow chart) performs the updating operation based on an event-data-file; this program generates an updated control file from the current control file. The event-data-file reflects the current updating information. That is, as information is received that results in a change of status (moving from one event point to another), an event-data-file record is generated and added to the file.

In conjunction with the event-data-file there is a table of endogenous information that may result in updating. This mode of updating applies to groups of observations in the file that have like characteristics where each is changed to a new status. In addition to possibly changing the status of groups of records, this table serves another function: to define the event structure for the survey in a machine readable format.

An additional element of updating exists whereby for each event two or more levels are defined with an associated time duration and disposition. When a participant has been at a given level for the allowed time, an internally generated event update occurs. Thus, following the expiration of an event level associated with the mailout of a survey questionnaire, for example, the ASCS automatically advances to the next level within the event and generates a report transaction. The transaction notifies the operational staff that a prompting postcard should be mailed to this individual. This report record is also acknowledged by the label producing software and a label record is produced for this postcard. All notices and label records associated with a given action are output for easy dissemination to the operations staff conducting the survey.

After a given sample participant's record has been updated, the ASCS accesses the next sequential record and continues until all records

have been processed. Next, summary reports are produced and a sort step, to group the individual diagnostics and notices for printing and label production, is entered. The final step is a run-terminating program which, if previous errors have not been detected, copies the backup file to the control file so that it and one backup are always identical.

The ASCS also updates the catalogues to reflect which files were used on the run and to turn off the active file indicator for the active event-data-save file. Since this final step is the only one that actually destroys any existing files, restarts of the ASCS can be made at any job step without having to have special restart procedures.

FEATURES OF THE ASCS

The ASCS has several user options that reflect the power of the system(3). These options give the user the ability to extend the system to meet his needs. Such things as defining reports or changing the event structure can easily be done through input controls. The following highlights the more important user options:

1. *Easy Addition or Deletion of Events Monitored by the System*--Each aspect of a survey operation is referred to as an event, such as the receipt of a survey questionnaire, or its failure to pass machine edit, or the mailout of a reminder postcard or thank you letter to a participant. These events are defined to the ASCS by a set of event description cards (one card per event). New events can be added or old ones deleted at run time by simply introducing or removing an appropriate event description card.

2. *Constant Monitoring of Each Sample Participant's Record with Associated Actions and Messages*--The ASCS monitors each participant's control file record, updating this record based on both externally and internally generated events. An example of an externally generated event is receipt of a response from a sample participant. Internal events are specified by the time allowed at each stage of the survey operation, making available to the user the equivalent of an automated suspense file for each survey activity.

3. *General Grouping Capability*--This feature allows a user to define and form subsets of the control file as a function of any variables contained in the file.

4. *Summary and Listing Reports*--Two types of reports are available, summary information and listings of the participants. One or both reports for each user defined group (as per option 2) may be requested at run time.

5. *Historical Information*--The ASCS provides a complete time sequenced history on all of the defined events occurring from initial mailout to edited, machine-readable data on each sample participant.

6. *Event Setting Capabilities*--Event indices can be set in participant records via control cards for all participants falling into user defined groups. This feature makes it easy to initialize the file to a specified set of events or introduce new event settings during the course of the survey.

7. *Simultaneous Event Monitoring*--The ASCS can monitor more than one event at a time for a given participant. Thus, two or more different types of data collection forms associated with the same individual can be followed simultaneously.

An important feature of these options is the flexibility they offer after the system is established and running. Many unforseen modifications can be made even when a procedural change occurs in a survey operation.

CONCLUSION

The system described in this paper represents an innovative approach to survey information delivery. The value of the ASCS can only be determined by evaluating the various components of automated survey management in comparison to the traditional manual control system. Major components of any type of survey control system that should be included in such an evaluation are: budget, project timetable, personnel availability, managerial control flexibility, and event or status reporting. Generally, once implemented the ASCS is a more efficient method of maintaining managerial control of survey activities. The ASCS provides operational control requiring minimal project staff attention thus resulting in a reduction of direct management labor and cost. A high degree of flexibility is afforded by the system in monitoring designated survey events and providing reports of past, as well as current, work status. Cross tabulations of various survey events can be performed, generating statistics not previously produced in event reporting. These characteristics are practically nonexistent in a manual control system because of the great amount of manual checking, collating, and verification required to accumulate and produce data.

Because survey events are updated as changes occur during data collection activities, ASCS reports are representative of the current status of field activities. This capability provides survey managers and planners with the ability to effectively allocate resources for future project needs as well as maintenance of budgetary control over current and projected survey activities.

When the ASCS is evaluated within the context of these components, it is evident that the system provides a greater capability for survey

management than has heretofore been realized by any other type of survey control process.

The utlimate benefits of the ASCS to survey research are numerous. Control over various survey activities is enhanced through the speed and efficiency of the ASCS's reporting and information delivery system. Also the effectiveness of allocating project funds and personnel to survey tasks when using ASCS was demonstrated by a reduction of the professional project staff required to manage the survey.

Another important result of the implementation of the ASCS is the day to day monitoring of data collection activities. Through daily monitoring of field events, project staff are informed of the current status of each sample unit as well as the project overall. Because the ASCS provides reports representing the current status of survey activities, data collection problems are easily and quickly identified thus facilitating immediate resolution of the problem situation.

Recommendations for use of the ASCS in future research projects are obvious. Whenever a survey incorporates into its operational structure many parts, with each part functioning as a separate survey event, then constant and accurate reporting of survey activities is tantamount to a successful operation. Fragmentation of a study into events is usually characteristic of large scale surveys but can occur in smaller ones as well. Therefore, regardless of the size of the survey operation, it is the responsibility of the researcher to evaluate the impact of the management system upon his and the client's needs as well as the fiscal resources available. If the needs are many, then the choice is made for the researcher: the Automated Survey Control System.

References

1. The ASCS was developed and implemented under Contract No. NO1-NS-4-2334, "National Study to Determine Incidence, Prevalence, and Cost of Head and Spinal Cord Injury", National Institutes of Health, DHEW.

2. Thornton, R.H., Brinkley, M.F., Wright, J.C. 1975. "An Automated Survey Support System", Research Triangle Instftute, 1975 July. The system described in this paper is the product of continued research in computer applications to survey management and therefore has drawn heavily upon the experiences of its predecessors. An earlier application of a support system was developed under contract (OEC-0-73-6666) to the National Center for Educational Statistics (NCES) of the Office of Education to support the National Longitudinal Study of the High School Class of 1972.

3. Brinkley, Margo. 1976. "Control System Documentation", Research Triangle Institute, 1976 August.

AN APPLICATION OF MARKET RESEARCH TECHNIQUES TO THE DISSEMINATION OF SCIENTIFIC AND TECHNICAL INFORMATION

Ethelyn F. Bishop and Audrey Clayton
Forecasting International, Ltd.
Arlington, Virginia

INTRODUCTION

Over the past several years, it has become increasingly apparent that effective dissemination of scientific and technical information can be considered a marketing problem. Practitioners of both information science and marketing are seeking to satisfy the needs of user groups ("consumers") through the development and distribution of appropriate products and services. Specifically, information needs of potential users of scientific and technical information (STI) -- research scientists, R&D and production personnel, teachers, and management personnel -- must be matched to actual and potential information sources. Users must then be made aware of the existence of the STI source. Finally, the source must be made available to the user in an easily accessible form at the time and place where it is needed.

To date, producers and distributors of STI have had insufficient information about the user and his needs on which to target their dissemination efforts. At the same time, because of the ever increasing quantity of STI available, it has become more and more difficult for the user to identify and extract the specific quantum of information which he needs for his current purpose.

Forecasting International, Ltd., under the sponsorship of the National Science Foundation's Division of Science Information is currently completing a study designed to provide information to producers and distributors of STI, and to information intermediaries, which will make efficient identification of user needs and delivery of appropriate information sources possible.

The project is actually developmental. Its purpose is to design and test a methodology for surveying the users of STI. The study uses the market research technique of "market segmentation" or cluster analysis to identify segments of the user market with similar STI needs, and to describe these groups in terms of their common STI use characteristics, their personal characteristics and their access requirements, thus allowing the development of strategies more responsive to those needs.

METHODOLOGY

Market segmentation is one of the major innovations to hit the marketing field in the past 20 years. In essence, the theory states that:

- The market is not composed of a single homogeneous group, but rather of a number of groups or segments, each of which is composed of a number of consumers displaying a series of common attributes.

- It is possible to split a market into its component segments, based on responses to a series of questions.

- The data describing the various segments can be used to more carefully target marketing or dissemination efforts to specific consumer audiences.

A variety of techniques are used to conduct market segmentation, however most methodologies follow the same general approach. First, an instrument is developed to collect attitudinal data, psychographics, demographics, etc. concerning behaviors and attitudes related not only to the major product area, but to other consumption patterns as well. A computerized cluster analysis routine is used to group respondents displaying similar characteristics. Demographic, psychographic or attitudinal profiles are developed for each identified group. The group profile data is then used to develop marketing strategies targeted to each segment.

A brief description of a segmentation study conducted by a leading beer brewing company will illustrate the application of this technique to marketing consumer goods.

For years the brewer had followed the traditional beer marketing approach -- marketing strategies assumed a homogeneous market. A recent market study indicated, however, that the company's promotional campaigns appealed primarily to women and to upper-income consumers, neither of which constitutes a heavy beer drinking market. A shift in advertising emphasis to appeal to younger, more active consumers resulted in an increased share of the market. In addition, the brewer began to target products for other market segments -- light beer for the calorie-conscious beer lover and the 7-oz. "pony bottle" for women and older consumers who prefer to quaff beer in less than 12-oz. lots. As a result of these marketing measures and administrative changes, the company has tripled sales since 1972.

This case study demonstrates most of the attributes of market segmentation which can be applied to the STI market - identification of user segments, tailoring of dissemination efforts to appeal to different groups, and packaging to meet the needs of specific segments.

The methodology adopted by FI followed this classic pattern. A questionnaire was developed to collect data on user STI needs, perceptions and behavior. This questionnaire was administered to 201 electrical engineers and 201 chemists in personal interviews carried out nationwide. Each respondent was asked to rank the following types of scientific and technical information in order of descending need from one to nine, where "1" identifies the most needed type:

- Information on results of new research in your field of specialization
- Information on results of previous research in your field of specialization
- Information on results of new research in fields related to your field of specialization
- Information on results of previous research in fields related to your field of specialization
- Information synthesizing available technology (e.g., survey and summary information)
- Handbook type information
- Critical review type information
- Catalog type information
- Bibliographic information

These rankings were submitted to a computerized statistical analysis procedure called the "Howard-Harris Clustering Program" which grouped respondents with similar rankings of information needs.

The rest of the data collected in the questionnaire was cross-tabulated to the segments to construct group profiles. The analysis of this data is now in process.

PRELIMINARY RESULTS

Enough analysis has been completed to show that the methodology is useful: the STI user universe can be grouped into segments based on needs and these segments have other common attributes which identify them for targeting the distribution and promotion efforts of STI producers and distributors.

Six segments have been identified within each discipline included in the sample. Each of these groups is being described by its other characteristics which would be of use to information producers, distributors and intermediaries.

For example, Segment 3 of the electrical engineer discipline contains 28% of the electrical engineer population. This is the group of engineers who are most interested and involved in research. The segment's three greatest STI needs are for information concerning new research in their own fields, previous research in their own fields, and new research in related fields. This information is used primarily to keep up-to-date on developments in the engineer's own field, to acquire ideas or stimulation for his work, and to find specific answers to questions related to his work. Because of this segment's research orientation, the project team has nicknamed it the "Researcher" segment. "Researchers," to an overwhelming degree, use professional journals to satisfy their major information needs. Conference proceedings are also used to a significant extent. The Research segment is the best educated and most heavily academic segment. Researchers do a great deal of writing and publishing, and belong to more professional organizations than any other segment.

On the other hand, Segment 4 of the electrical engineers is involved in much more operational work activities, and the information needs of the segment reflect this less theoretical orientation. This group gave handbook and catalog type information the highest priority of any electrical engineering segment. The study team labeled this the "Working Engineers" or "Worker" segment. The "Worker" group, comprising 17% of the electrical engineer population, is heavily industrial and less well-educated than the other segments. "Workers" do not attend as many professional meetings or join as many professional or multidisciplinary societies as other segments. Also, they do not publish much and appear to be Junior Engineers and Instructors.

The "Worker Segment" uses free periodicals, professional and trade journals, catalogs and handbooks to satisfy major information needs. The information gained from these sources is used to find answers to specific questions relating to the engineers' work and to identify new materials, methods and procedures.

Further analysis will augment these descriptions to include users' perceptions of current STI sources, and user experience with channels of distribution and means of promotion.

In addition to these descriptions of the basic user groups, many subsidiary findings are emerging from the segmentation analysis concerning "most needed" sources of STI. For example, it appears that segmentation by STI needs may correlate more strongly with principal work activity than with discipline. In choosing chemists and engineers as the subjects of the demonstration survey, the team assumed a natural segmentation of the sample by electrical engineering vs. chemistry. However, during the segmentation analysis, "Researcher", "Worker" and "Developer" segments become apparent in both disciplines.

Professional journals are the most popular source for satisfying the prime information needs of electrical engineers and chemists as a whole. On the other end of the spectrum, current awareness services, data base access services, fact finding and reprint services are seldom used to satisfy top priority information needs. During ensuing analysis, the project team will attempt to determine if this lack of use is due to lack of awareness on the part of the users, poor accessibility of the source, or some other factor.

It also appears that the surest way of contacting electrical engineers to tell them of potential information sources is through their professional societies. Membership in professional societies of a disciplinary nature is very high. Further analysis of the data base will investigate the respondents' perceptions of the professional journals published by these societies as a means of promoting awareness of STI products and services.

CONCLUSION

This developmental study has produced a rich data base of information for use in planning STI dissemination to chemists and electrical engineers. Final results should be available by August 1977.

A MODEL TO PREDICT GROUPS MOST LIKELY TO PAY FOR "INFORMATION."

Lea M. Bohnert
Graduate Library School
University of Rhode Island
Kingston, Rhode Island

It must be recognized that "information" is a service, not a commodity. It is a service performed by personnel trained in interpreting requests for information and the recorded information, or documents, that may match the request. The request may be precise and easy to match with a well-known document or the request may be general, even ambiguous, and difficult or impossible to match with known documents. It is the document requesters who initiate the communication process known as retrieval. It is the job of the retrieval experts to deliver whenever and however possible a copy of appropriate documents to the requester. The traditional name for the retrieval expert is "librarian."

Secondly, it is important to realize that everyone pays for information, whether they know it or not. Most payments are delayed and indirect. Information about a thing, and the thing itself, often go together inextricably. Even if groups most likely to pay do not, or can not, pay, it is a useful exercise to identify groups requiring different kinds of retrieval. It will be shown that closeness between what is requested and what is available is an important characteristic of retrieval service that can be costed.

Two related theoretical developments will be employed as guidelines for the analysis. The first is the Mooers/Bohnert retrieval model (1), which displays the kinds of communication that occur between document producers and document requesters as mediated by librarians. The model displays the two main kinds of messages, messages about documents and the documents themselves.

Most importantly, the Mooers/Bohnert model shows that there are two kinds of communication which are separate but equally important to the regular work of librarians. One type of communication is initiated and maintained by the producers and distributors of documents. The purpose is to influence librarians to select and to acquire copies of documents in advance of specific document requests. The other type of communication is initiated by requesters. The purpose is to secure access to requested documents, if such documents are known to exist and copies are available.

The second guideline is the Notification Theory of Robert A. Fairthorne (2), in which six elements are determined to be essential in any systematic consideration of the retrieval process. The six

elements are Source, Code, Message, Channel, Destination and Designation. The first five elements are familiar from the Shannon model for general communication (3). The sixth element must be added whenever the meaning of the Message is significant as it is for document requests based on subject.

Mr. Fairthorne also demonstrated that a minimum of three elements, or a "triad", have to be considered in analyzing any retrieval situation. Two elements do not model a resolvable situation because too many alternatives remain equally feasible. To illustrate, it might seem obvious that the groups most likely to pay for retrieval service would be those who have a strong need for a requested document (perhaps an important decision will be affected) and those who are able to pay for the service. However, these conditions are necessary but not sufficient. Perhaps the kind of document requested does not exist or the document may exist but not in a form intelligible or convenient for the requester.

Of the six elements in the Notification, or retrieval process, the first and last are customarily outside the control of the retrieval service. Both the producers of the documents, or the Sources of the Messages, and the requesters, or the Destinations, decide what to produce or to request, and also, when and how to do so. Librarians endeavour to match, to identify, the documents requested among the documents known to be available.

Sometimes librarians may intervene by recommending that particular documents should be produced or by providing copies of appropriate documents to requesters before any request has been made. These "initiatives" are necessarily based on careful study of previous successful and unsuccessful matches of requested documents with known and available documents.

The best way to show the application of these theories is to present two different cases showing the extreme kinds of retrieval service possible. Complexity or difficulty of retrieval will be shown to depend on how closely the requested document matches documents known and available. In this discussion, "complexity" does not refer to the sophistication of the telecommunication or computer devices used by any particular retrieval service.

The first example will be the retrieval of telephone numbers that are expected to exist by the requester. The second example will be the retrieval of scientific and technical documents that are not expected to exist by the requester, such as a patent applicant.

In the first example, concerning telephone numbers:

1. The Source is authoritative andin the USA single (Ma Bell).
2. The Code is numeric and terse.
3. The Channels have been restricted by the Source to two choices:

a. "Directory assistance" over the telephone.
b. Directories in printed form.

Therefore:

4. The Messages that may be requested are severely limited.
5. The Designations possible to assign to the Messages are also limited.
   a. Best method of retrieval is by correctly spelled names of individuals or organizations.
   b. Less efficient are the "Yellow Pages," the subject classified directories.

So, the retrieval required in this example is simple because the service providing the Messages is the same organization that produced the Messages (assigned the telephone numbers). In this example, control by Source of the retrieval process is at a maximum.

In the second example, concerning patents:

1. The Source of the scientific and technical documents are expert and many.
2. The Code, or language, of the document is precise compared to everyday language.
3. The Destination, or requester, of the documents or Messages does not expect, even hopes if he is a patent applicant, that such documents exist.

Therefore:

4. The Channels are most diverse, published and unpublished, and in a variety of media and format.
5. The Messages are most various and complex.
6. The Designations appropriate to assign the Messages are also many and various. The U.S. Patent Office Classification System demonstrates the complexity of the subjects.
7. Finally, the Destination, the requester and the patent examiner, must spend considerable time reading possibly appropriate documents.

The resulting retrieval service is time consuming and complex. There is no central control over the documents being produced or the kinds of Designations appropriate to assign to the Messages. It is the most difficult kind of retrieval search that can be made, namely, for documents that an expert hopes have not yet been produced in his own field of expertise. The U.S. Patent Office performs daily this "impossible" kind of search and has established an elaborate appeals procedure to correct the inevitable mistakes.

The considerations presented do not exhaust the analysis of a particular retrieval situation but should indicate the main lines that must be followed in order to accomplish such an analysis.

An entirely separate question which is not treated in this paper is whether groups identified as able to pay for a suitable retrieval service should be required or allowed to pay because of other social values. One kind of retrieval in which retrieval might be diastrous is telephone access to poison information centers.

References

1. Bohnert, Lea M. 1975. "Neglected Definition of Information Retrieval." In: Information Revolution: Proceedings of the American Society for Information Science, Vol. 12, Washington, D.C., American Society for Information Science, 1975, 7-8.
2. Bohnert, Lea M., 1974, "Fairthorne's Triads as an Aid in Teaching Information Science." Journal of Documentation. 1974 June; 30(2): 210-215.
3. Shannon, Claude E.; Weaver, Warren. 1949. The Mathematical Theory of Communication. Urbana, Illinois: The University of Illinois Press; 1949.

THE VALUE OF ENERGY INFORMATION

James D. Cape
ERDA Technical Information Center
Oak Ridge, TN

I am going to speak about the value of energy information in the context of change of ERDA's technical information program. I am going to speak about the closed circle of technical information on nuclear science and engineering and the open circle of technical information on all energy. By we, I refer to the technical information program of the AEC and its successor, the U.S. Energy Research and Development Administration.

Information in the federal government is considered to be a public good. The value is understood and not questioned, other than the size of a government agency's technical information program. The closed circle of technical information on nuclear science and engineering comes about by information generated by scientists and engineers for other scientists and engineers. The circle was closed since scientists create information to be used primarily by other scientists. Progress is made by building on the work of others, knowledge begets more knowledge. There were offshoots from the circle, e.g., to students to encourage their scientific and technical interests and create more scientists and engineers, and the general public to increase general understanding of technology, and through understanding--acceptance. The accepted goal of advancing technology and its application was not questioned. The public was to be "educated" and was not part of the circle. In retrospect this was a mistake, however I don't believe that it was an avoidable mistake but intrinsic to any high technology. High technology information has value to those interested in high technology and who can use this information. Low technology information has value to those interested in low technology and who can use this information.

Energy information is both high and low technology. High technology can be understood and used by only a few. Low technology can be understood and used by many. Information has value only when applied to a need. To apply information, the potential user must learn that particular information exists for which the user has a need, and the user must obtain this information.

To facilitate this, the ERDA Technical Information Center obtains the world's literature on energy, creates a computerized data base of bibliographic citations, subject indexing, and abstracts describing this literature from which a number of information products are produced. An important type of information product are abstract journals: ERDA Energy Research Abstracts announces all literature resulting from ERDA support for those who believe that value of government R&D support is its publications; Energy Abstracts for Policy Analysis for those interested in the managerial and analytical literature; monthly update journals for solar, fossil, and geothermal announcing all literature added to our data base in one month's time for those finding value only in these areas; and Power Reactor Docket Information for those interested in only documents supporting the licensing

of commercial nuclear reactors. Bibliographies are also published from our data base for those finding value in evaluating the complete literature in a particular area. These publications are of value to those who want to know what information is available and how to get it. The data base is used for other purposes as well, including use by outside ERDA organizations offering SDI and retrospective searches. What is the value of these information products? The value to ERDA is avoiding duplicating research and aiding development. The value to the taxpayer is making the tax dollar spent by ERDA more effective.

Abstract journals, bibliographies, magnetic tapes, SDI and retrospective searches have value mainly to those in the closed circle who are involved with high technology. The challenge to the ERDA technical information program is to break out of this closed circle and provide technical information of value to individuals, commercial firms, small industrial firms, architects and engineers, etc., who will apply this information, and not necessarily increase technology. The closed circle has opened to include a large number of not easily defined information users. These users must be reached with information products and services, and some sort of feedback must be obtained as an indication of value to affect plans for further information products and services. Much of this can be accomplsihed by doing what we have been doing but on a large scale. However, new programs are being instituted to determine what information products and services are of value.

There recently has been a shift of emphasis of working on long-range solutions to energy shortages to near-term approaches. In other words, information on comparable low technology conservation and solar heating for industry and individual homes is being emphasized. The problem is what type of information in what form will be of most value to the user, and how is this determined.

One way of determining what information is of value to the public is developing an information product and offering it to the public. This ERDA did with ERDA-75 "Catalogue of Solar Energy Heating and Cooling Products". The response was more than expected. We received and filled requests for 10,000 copies and referred further requests to GPO. GPO depleted their stocks and ERDA has again reprinted this publication. Obviously, this information product is of value.

Another method determining value is sending out notices to mailing lists asking for replies to those wishing specific publications or those wishing to be put on mailing lists to receive publications or notices on a regular basis. The receiptor must make a decision that the information offered is of value to the receiptor and must take some action to receive this valued publication. ERDA programs have followed this method to develop mailing lists for their publications and notices.

In the first five days of March 1977, the ERDA Technical Information Center sent out 507,000 letters from Dr. James Schlesinger asking for national energy policy recommendations. Over 10,000 recipients of that letter took the time and effort to reply. The problem with this approach is how to evaluate such a large number of replies. Each reply was read

and evaluated by a trained evaluator. The opinions expressed in each reply were translated by the evaluator into answers to predetermined questions, and these, together with up to 100 characters of free text expressing opinions not fitting the predetermined questions, were fed into our computer. Our computer provided the totals for the predetermined questions and listed the free text opinions, and these were forwarded to the White House. I mention this project as an example of how we could use mailing lists to ask for opinions on what energy information is of value. The expense of reaching a good sized sample of the users of energy information and evaluating returns is substantial.

Another approach to the problem of determining the value of energy information to the user is the Energy Extension Service (EES). The EES will work through state Energy Offices and universities to develop a county agent network working with individual companies, architects, manufacturers, associations, and homeowners to provide advice and publications of value to the ultimate user of this information. An important aspect of the EES is feedback to ERDA of what information products are of value to the information user. The ERDA Technical Information Center will provide searching of its data base and publications. The value of these will be determined by feedback from the users of these services and products.

Feedback is all important with energy information because the technology must be accepted and applied widely as a result of an evaluation of benefits and costs (economical, social, and environmental). Value can no longer be understood and untested, but must be measured by the user and transmitted back to the information product and service provider.

In this paper I have outlined the change in our technical information program resulting from a different type of "customer" requiring a different, more direct, method of determining value.

A COST/BENEFIT EVALUATION OF ONLINE INTERACTIVE BIBLIOGRAPHIC SEARCHING IN A RESEARCH AND ENGINEERING ORGANIZATION

A. D. Collette and J. A. Price
Exxon Research and Engineering Co.
Florham Park, N.J. 07932

Summary This paper describes the experience since early 1974 of computer-assisted online, interactive bibliographic searching at the Florham Park, N.J. facility of Exxon Research and Engineering Company. It particularly focuses on the results of a survey conducted in the fall of 1975 designed to quantify the real value of the searching service to the engineering and chemicals population. Obviously, some value was anticipated since users showed a continued willingness to pay for the service, contributing to its growth during the period.

The analysis presented here has been comprehensive and detailed. Strict attention has been given to an accounting of all the costs associated with carrying on the searching activity in addition to the computer connect time charges. Included in costs are the information searcher's time at the terminal plus time required for planning strategy, statistical analysis and customer discussion. Additional costs are those for terminal rentals, maintenance fees, space charges, offline print charges and time spent by the technical person in assisting the search. Benefits have been obtained through the use of a questionnaire completed by users of the service. Quantification of the benefits was derived from estimates of time saved by the technical person plus dollar savings to the individual himself or to the corporation more broadly. Precedence for deriving benefits in this manner has come from an as yet unpublished in-house study on the value of the external literature to research personnel in our company.

The principal result of this approach has been to yield a value in terms of benefit-to-cost ratio of nearly three to one. Further, other statistics shown an average search in this survey has taken 49 minutes computer time and has utilized 3.2 data bases.* The principal conclusion is that even though the costs to produce a completed search are unprecedentedly large in comparison to other literature figures, the benefits far outweigh the costs confirming that the searching activity has a payout comparable to or greater than other corporate operations. Thus, while a larger b/c could have had more impact on management and the information profession, we have strived to be as conservative and realistic as possible. Further, we believe no other published study has been as thorough in developing costs nor as exacting in identifying and quantifying the resultant benefits to customers.

---

*Even though overall searching in 1975-7 has required substantially less time and fewer data bases, we have of necessity retained in this paper the original figures. The reader is reminded that the survey was conducted during a period of rapid proliferation of new data bases which our searchers were investigating for relevance. Moreover, the searches included in the survey were longer than the normal average for 1975--14 minutes per data base--because of the intended emphasis on substantial searches. Current search time per data base is running about 12 minutes.

Introduction In recent years it has been popular to ascribe the words "management" or "managing" to various aspects and activities of information centers and special libraries. Sometimes this relates to the running or administering the information function--witness articles or symposia like: How to Manage and Use Technical Information,(1) "The Emerging Information Manager: Bridging the Gap,"(2) "Technology, Management and Economics of Information Centers and Services,"(3) "Innovative Management of Technical Information Functions in Industry."(4) Sometimes the orientation is more related to what the information function can do for management. Literature examples are "The Manager and STI,"(5) "Special Libraries: A Guide for Management,"(6) "The Information Center: Management's Hidden Asset."(7) At other times the management connection with the information function is in the role of examining the effectiveness and costs of operating the function. Typical examples of this feature are "Managing the Costs of Information,"(8) "Cost/Benefit Analysis of Company Information Needs,"(9) and "How to Survive in Industry, Cost Justifying Library Services."(10)

The present study is closely aligned with this latter aspect. It zeroes in on the value--cost effectiveness or benefit-to-cost ratio--of one information activity, namely, online interactive bibliographic searching. It was not expected that this would be an easy task since the literature indicated certain problems with accomplishing this. One source(9) indicated that "there is seldom any clearly thought-out framework for formal evaluation of net values"; another(6) said "even though it is possible to determine the dollar costs of specific library activities, it is difficult to determine the dollar benefits." A third source(11) concluded that "it is clear that information is not susceptible to the usual attempts at cost-benefit analysis...". However, we had just completed an in-house assessment of the value of external literature to research personnel and felt assured that the same techniques and principles were applicable to the case for online searching. An IBM study also recently completed gave credible b/c results.

Thus, while there appeared to be a lack of relevant published information on the value of interactive computer searching, there was a distinct need to obtain it. The approach used and the findings are the subjects of this paper.

Background Online bibliographic searching has been employed for the Florham Park-based engineering and chemical staff since April 1974. Prior to this date, external searches were done on a more limited basis using the API literature and patent data bases. From 1965-70 these were batch-mode computer-searched using a Fortran program which was cumbersome and costly to operate. About 1970 these data bases were re-programmed for API on the [1]INQUIRE system--a purchased program--which was also searched batch-mode. Subsequently, due in part to the high-cost of disk pack storage, we contracted our searching of the API files with their subscriber service in New York City, paying a rate per batch-mode search of sixty-five dollars.

For searches on data bases other than API during this period, we used institutional or government information center facilities. Early in 1974 we contracted with [2]SDC for online access to the API data bases. At the same

1 - Product of Infodata Systems of Falls Church, VA.
2 - Systems Development Corp., Santa Monica, CA.

time SDC provided us ready access to additional data bases such as Compendex (Engineering Index), Chemcon (Chemical Abstracts Condensates), NTIS (National Technical Information Service), and others. Online access greatly expanded the scope of our capabilities and provided rapid service for our clients. Later in the year we added [3]Lockheed's service giving us a much larger base for bibliographic searching. Growth in usage has been steady, but planned (see below). At the present time, we conduct about fifty searches per month utilizing up to twenty different data bases.

In order to promote usage of our online retrieval system we prepared two announcements of its availability. One was directed to the engineers in the technical area and the other to our corporate and management professionals. We invited both groups either to take part in searches, or to present their requests to be handled at our convenience and discretion.

Moreover, in order to stimulate greater interest in online searching and its potential we arranged demonstrations for small groups of engineers with specific interests and prepared searches relevant to their interests. Later we were asked to demonstrate the technique and capabilities for a management committee. Due to the favorable response of this group, a request was made for a large scale demonstration for about 200 professionals. This was accomplished by using an online terminal connected electronically to a video amplifier/projector so that the entire group could see the "live" demonstration on a large screen in an auditorium. Several searches were shown while we gave an oral description and explanation of the procedure. The overall effect of these promotions, as anticipated, was responsible for a large increase in demand for searches by former customers as well as new users of the service.

The search team engaged in this work is comprised of two persons--one with extensive experience and background in engineering files and indexing systems; the other with background in analytical petroleum research and in traditional information searching. Besides searching they have been responsible for maintaining an in-house computer storage and retrieval facility consisting of a fifteen year accumulation of proprietary documents and surrogates. These have all been coordinately indexed and currently are batch-mode retrievable by the INQUIRE software program. These two persons are thus quite aware of the problems associated with the building and maintaining of data bases.

Discussion During the growth and development of online searching we have kept detailed records on data base usage by vendor, by data base, by company department use, and by the number of data bases used per search. These data allowed for calculation of the average number of data bases accessed per search and the average time per access, which show 3.2 accesses per search and 16 minutes per access. Our figures coincide with data reported earlier by Calspan(12) in a questionnaire survey on online searching.

While gathering this statistical data we realized that a stewardship report on the costs and benefits of the online searching function was needed. It was necessary from the point of view of properly managing future activity and growth. But it was also necessary for demonstrating to corporate management the utility and value of the activity.

---

3 - Lockheed Information Systems, Palo Alto, CA.

At this time we speculated on how best to obtain feedback from clients on the usefulness of the information supplied to them--in particular, we wanted to know to what extent the retrieved data benefited their engineering projects. It was finally determined that a questionnaire would be the best mechanism for getting the required answers. Since a questionnaire is sometimes viewed as a nuisance item, we decided to make it as brief and simple as possible. We limited it to three questions, and elected to keep questionnaires sent to repeat users to a minimum. This also helped to broaden the user base for the evaluation.

When sending out questionnaires it was decided to send to those people who had submitted substantive search requests rather than to those who simply wanted to locate a specific citation by author or organizational source. As noted elsewhere, this has the effect of lengthening the average time calculated for searching because requests for specific items usually result in more rapid searches.

In order to allow sufficient time for the recipients of a search to use the information supplied to them and to realize some measure of benefit from it, the questionnaire distribution lagged the actual request by about three months. This was considered as a compromise since more benefit may have resulted from more elapsed time. On the other hand, the identification of benefits resulting from the search may have been less precise at a longer period of time. In the survey of the effect of the external literature on research personnel referred to earlier, the benefits were identified and measured immediately at the time of receipt of the information. So this lapse time feature may not have been necessary or critical.

The actual questionnaire used is given in Appendix I, and the actual responses for each month--September and October 1975--are summarized in Appendix Tables II and III. For purposes of the present evaluation, the results from September and October have been combined. In actual fact the survey questionnaire was changed slightly in October (1) to include a dollar category lacking in the September version, and (2) by transposing questions 2 and 3 to give more importance to the time-saved question. Nevertheless, in essence the assessment covers the two months as one period of time.

A summary of the combined responses to the survey is given in Table 1.

Table 1
ONLINE SEARCH SURVEY RESPONSE

Total searches during period	86
Questionnaires sent out	59
Questionnaires returned	46 (78%)

The recovery rate of 78% is considered to be excellent and in part reflects the interest and value users place on online searching.

Table 2 shows the number of responses to each of the questions--also in combined form.

Table 2
NUMBER OF RESPONSES (%)

(1) Was the search useful--either in providing information or in indicating that there wasn't any? YES 39 (85) NO 7 (15)

(2) Please estimate your time saved:

Less than 5 hours	14
5-20 hours	23
More than 20 hours	7
No answer	2

(3) Please estimate dollar value over and above your time saved:

Less than 50	16
50-100	10
100-500*	8
More than 500	8
No answer	4

*This category not provided in September questionnaire.

It is striking that seven (15%) people said that the information unearthed by the search was not useful to them. Most of these respondents checked the lowest categories of benefit as expected. At first, it seemed anomalous to allow any quantitative benefit to accrue to these respondents. However, even though the information retrieved provided no apparent benefit, these users were indicating that the information searches actually reached this point faster than the user could have done--thus effecting a corporate savings in time. One company survey found that "for every hour the library (people) spent literature searching, we saved the engineering staff 8.86 man-hours of effort."[10] The review publication[6] commented that "these data are based on estimates by the users of the time they would have needed to accomplish the results provided by the library, so they must be viewed as an indication, not an absolute measure." It is clear that in the present study users were not considering the magnitude of benefit to be this large. Also, since we had not emphasized in the questionnaire the probable faster search time, we doubt that the user's estimates included this aspect.

In evaluating questions 2 and 3 in Table 2 for calculations of benefits or savings, we used the mid-point or low end of the range. For example, in question 2 the '5-20' hour range was taken as 10 hours; in question 3, 'more than $500' was taken as $500. For this reason, we believe that the estimates for benefit finally developed are ultraconservative. There is little doubt that the benefit could have "ballooned" by using the top numbers in each case. The distribution of answers in both questions 2 and 3 indicates that the majority of users had beneficial impacts in lower categories than the top. We thus feel there is no exaggerated skew to the data.

Using the numbers from the table as indicated, the estimate of time saved by users came to a total of 433 hours. Using a literature[13] figure of $22,000 in 1975 for the median annual salary of AIChE members and a 40%[14] increment for petroleum industry fringe benefits such as health care, vacation

plans and the like resulted in an hourly rate of about $17.00. This also happens to be close to the hourly charges listed in an article titled "The Information Brokers"(15) for searching services. Converting the hours gave a quantified benefit amounting to $7,360. From question 3 the estimate of dollar benefits amounted to $7,150. The combined benefit totaled $14,510.

For 46 searches, this amounts to a benefit of $315 per search. While this is a substantial value, the cost of providing this benefit needs to be factored in and our approach is given below.

Traditionally in this field the computer connect time has been used as the dominant cost, and in some cases the only cost. In fact, it has been touted that "the era of the $5 search" has arrived. We believe that this kind of salesmanship is harmful and creates a credibility problem for the more reasoned sector of the online community. Our repeated point is that a realistic approach to developing costs still shows that online searching is a cost-effective service to technical and management personnel.

In this study, if we were ultraconservative in calculating benefits, we were ultraliberal in developing costs. The indirect as well as direct charges associated with operating the searching function were cranked into the total cost. For purposes of the calculation, we have used the computer time for our location in New Jersey @ $1.10/minute which included the Tymshare cost in 1975 of $10.00/hr. For the data bases used most during the period of the survey, the costs were $45-$65/hr.

Table 3
COST ELEMENTS FOR SEARCHING ACTIVITY

- Computer connect time @ $1.10/minute (incl. Tymshare)
- Offline printing (10% of connect time)
- Searchers' time (68 hours; about 38 hours at terminal)
- Engineers' time (10 minutes/search)
- Lease cost including maintenance (~$100/month)
- Space charges and miscellaneous overheads

The costs for offline printing have run between 8-13% of the computer cost. So for the cost calculation, 10% was used.

The searcher's time charges are a significant portion of the overall cost. The time spent on a search is not just the time interacting with the computer but also the time spent in developing strategy and in discussing the search question with the technical person. Other searcher costs which must be considered are those connected with keeping statistical analyses and with maintaining searching skills such as attending seminars and meetings, doing background reading and so on. In short, the total overhead charges for supporting the searchers in their work must be included in the calculation. To conduct these searches required 68 hours of searcher effort; actual time at the computer terminal was about 38 hours.

For each search some time on the part of the user is required. Sometimes it is only time needed to structure the question; sometimes it includes assisting the searcher at the terminal. For purposes of this study we estimated an average engineer spent about 10 minutes/search.

The cost of renting and maintaining computer peripherals amounted to ~$100/month. This is a pro-rated figure based on the number of searches in the study vs. the number actually performed during the period. In cases where equipment has been purchased, a depreciation figure should be used.

Space charges for equipment and personnel need also be considered and included. So do materials costs, administrative overhead and the like.

The overall costs for the 46 searches when cumulated this way amount to $5,132. Thus, each search costs about $112.00, which is one of the largest--if not the largest--amounts reported in the literature.

Comparisons of the total cost of the total benefit can be done in at least two ways. The net benefit is the total benefit less the cost. So in this study the net benefit = $315 - $112 = $203, which is the savings resulting from each search conducted. The more conventional method used by value estimators is the benefit-to-cost ratio. In this case b/c = $315/$112 = 2.8/1.

Conclusion Value determinations such as cost-benefit studies on the worth of information services are clearly lacking in the information literature. The reasons for this stem from beliefs that it is impossible or infeasible to quantify these kinds of activities, and from beliefs that it is not necessary because "it is obvious that information is invaluable." Yet in these days of the "bottom line" fever there are pressures from both internal and external management to examine information services on the same basis as other corporate operations.

This study presents a conservative but practical approach to developing hard costs and benefits for online bibliographic searching. We believe this evaluation has solidly established the worth of this service. It is hoped that others will develop case studies of their own using similar techniques and guides.

REFERENCES

(1) Dyke Jr., Freeman H., How to Manage and Use Technical Information, Industrial Education Institute, Boston, Massachusetts, 1968.

(2) Information Industry Association 1st Annual National Information Conference and Exposition, "The Emerging Information Manager: Bridging The Gap," April 18-21, 1977, Washington, D.C.

(3) Engineering Foundation Conference, "Technology, Management and Economics of Information Centers and Services," August 29-September 3, 1976, Easton, Maryland.

(4) Engineering Foundation Conference, "Innovative Management of Technical Information Functions in Industry," August 14-19, 1977, Rindge, New Hampshire.

(5) Wolek, F. W., "The Manager and STI," prepared for IRI Task Force on STI, October 1976.

(6) Special Libraries Association, Special Libraries: A Guide for Management, New York 1975.

(7) Meltzer, Morton F., The Information Center: Management's Hidden Asset, American Management Association, New York 1967.

(8) Strassman, Paul A., "Managing the Costs of Information," Harvard Business Review, September-October, 1976, pgs. 133-142.

(9) Wills, Gordon and Martin Cristopher, "Cost/Benefit Analysis of Company Information Needs," UNESCO Bulletin for Libraries 24 (No. 1): 16 (Jan.-Feb. 1970).

(10) Kramer, Joseph, "How to Survive in Industry, Cost Justifying Library Services," Special Libraries 62 (No. 11): 487-489 (Nov. 1971).

(11) McMullen, R. M., "Information - Value and Cost," Letter to the Editor, Journal of the American Society for Information Science, September-October, 1973, pg. 404.

(12) CALSPAN Corp. Questionnaire Survey, October 25, 1974.

(13) Chemical and Engineering News, October 27, 1975, pg. 41.

(14) Mobil Co. house organ., December 1976.

(15) "The Information Brokers," Bulletin of the American Society for Information Science, Vol. 2, No. 7 (February) 1976.

ENGINEERING INFORMATION CENTER

COMPUTER SEARCHING

DATE_____

TO_____

ROOM NO._____

We are continuing to evaluate our on-line searching services. A recent survey covering the month of September, 1975, has been completed. The results indicate further periodic evaluations are justified.

Thus we would appreciate your cooperation in providing information on the results of another sample of searching. This survey will cover searches done in October, 1975.

We would like to establish a dollar value, if possible.

This information will be helpful in future planning here at the E.I.C. and will also be used to compile data on the value of information retrieval for the EJC (Engineers Joint Council) Commission on Engineering Information.

Search No. 75- Date:

Subject:

(1) Was the search useful - either in providing information or in indicating that there wasn't any? YES NO

(2) Please estimate your time saved:

Less than 5 hours _____

5 - 20 hours _____

More than 20 hours _____

(3) Please estimate dollar value to Exxon over and above your time saved:

Less than 50 _____

50 - 100 _____

100 - 500 _____

More than 500 _____

(If you can be more specific about either time saved or value to Exxon, please insert your own numbers.)

COMMENTS OR SUGGESTIONS:

Please return to:

APPENDIX TABLE II

SEARCHES - SEPTEMBER 1975

Search Number	Computer (Minutes)	No. of Data Bases Accessed	Average Time of Access (Min.)	Searches Time (Hrs.)	Questions 1	Questions 2($)	Questions 3(hr)	Remarks
75-312	105	5	21	3	Yes	>500	5-20	
75-313	52	3	13	1.5	Yes	<50	5-20	
75-315	73	3	24	2.5	Yes	>500	>20	
75-316	16	1	16	0.5	Yes	>500	5-20	Favorable comment
75-318	28	3	9	1.0	Yes	>500	5-20	"Exposed coverage"
75-320	53	3	13	1.5	Yes	50-100	<5	Favorable comment
75-321	38	4	9	1.0	Yes	50-100	5-20	
75-324	39	4	9	1.0	Yes	<50	<5	
75-326	73	4	18	2.0	Yes	50-100	5-20	"Expand coverage"
75-331	18	2	9	0.5	Yes	50-100	>20	"Excellent System Service"
75-335	64	4	16	2.0	Yes	<50	<5	"Definite value, but cannot evaluate"
75-336	50	2	25	1.0	Yes	50-100	<5	
75-338	33	4	8	1.0	Yes	>500	>20	
75-340	48	4	12	1.5	No	-	-	
75-341	17	3	5	1.0	No	<50	<5	
75-347	34	2	17	1.0	No	<50	<5	
75-353	42	3	14	2.0	Yes	50-100	5-20	
75-355	102	5	20	3.0	Yes	>500	>20	"Job well done and very useful'
75-356	44	4	11	1.0	Yes	-	>20	
75-357	25	5	5	0.5	Yes	50-100	<5	
75-358	50	5	10	1.0	No	<50	5-20	No references "good news"
75-360	35	5	7	2.0	Yes	<50	<5	"Not too successful"

APPENDIX TABLE III

SEARCHES - OCTOBER 1975

Search Number	Computer (Minutes)	No. of Data Bases Accessed	Average Time of Access (Min.)	Searches Time (Hrs.)	Questions 1	Questions 2(hr)	Questions 3($)	Remarks
75-362	45	2	23	2	Yes	5-20	100-500	"Speed and cooperation outstanding, valuable service"
75-363	43	7	6	2	Yes	<5	<50	
75-364	29	2	14	1	Yes	5-20	100-500	
75-366	8	1	8	0.5	Yes	5-20	100-500	
75-369	56	3	18	2	Yes	5-20	100-500	
75-371	16	2	8	1	Yes	-	-	
75-372	100	4	25	3	No	>20	100-500	"Too many references, mostly useless, etc."
75-376	64	4	16	2	Yes	10	-	
75-377	33	1	33	1	Yes	5-20	50-100	"Search substantiated a hunch"
75-378	27	6	4	1.5	Yes	<5	50-100	
75-379	21	4	5	1.5	Yes	5-20	<50	
75-380	90	2	45	2.0	Yes	5-20	100-500	
75-381	68	4	17	1.5	Yes	5-20	100-500	
75-382	44	3	15	1.0	No	<5	<50	
75-383	24	1	24	1.0	Yes	<5	<50	
75-384	12	1	12	0.5	Yes	5-20	>500	
75-385	94	3	31	2.0	Yes	35-40	>500	
75-387	98	2	49	2.0	Yes	<5	<50	
75-388	43	2	21	1.0	Yes	5-20	<50	
75-390	89	3	27	1.5	Yes	10	320	
75-391	83	4	20	2.0	Yes	5-20	<50	
75-393	53	3	17	1.5	Yes	<5	<50	
75-395	65	2	33	1.5	Yes	<5	<50	
75-396	32	3	10	1.5	Yes	<5	50-100	

52

55

THE IMPORTANCE OF EFTS IN THE DEVELOPMENT OF INFORMATION SYSTEMS FOR THE GENERAL PUBLIC

Arthur B. Corte
Department of State
Washington, D. C. 20520

If information has value, as the theme of this conference implies, the information oriented societies envisioned for the future will require a mechanism for the receivers of information to exchange something of value with the suppliers; i.e., to pay for it. We accept this concept when we pay a lawyer or consultant for advice, however, when the information comes in print form, the existence of libraries and advertising hides the cost of acquiring information from books and magazines. The development of computer and telecommunications based systems for disseminating information will focus attention on the need to pay for it and convenient means for doing so will be necessary for such systems to fully develop. This will especially be the case for systems oriented to consumer needs where the universe of potential users is in the hundreds of millions.

Computer based information systems serving institutions are commercially operational today, consumer oriented systems should soon follow. The home computer terminal is already here in the guise of television games like "Pong", the addition of an alphanumeric display, a keyboard and enough memory and processing capability to permit interaction with a computer will come about in a few years. While such terminals will find some use working with home computers for various purposes, its most useful function will likely be to display information obtained from suppliers scattered around the country, and hopefully, the world.

Much of the existing distribution system for information intended for the mass public is inefficient and unrewarding to the information generator, most of the system's costs are involved with the production and distribution of the paper medium. After deducting the retailer's 40% markup, the costs of typesetting, printing, binding, warehousing and transportation, the residue left to pay for the authoring and editing functions that comprise the generation of information is a small part of the purchase price of a book. The same principle holds more or less true for magazines, depending on how much advertising support they attract.

Information distribution systems based on home terminals, computer storage and telecommunications do

away with the high production and distribution costs of paper based media and bring the cost of "publishing", in the sense of making available to the public, so low as to permit universal self publishing and making the price of information to the user that which the author wishes to charge for it plus a small distribution charge to cover mechanical costs.

The technology for such services is easily at hand. The simplest and cheapest form of electronic publishing could be a variant of the telephone recording, delivering, instead of a recorded voice, the analogue equivalent of a string of characters that will be displayed upon the screen of the home terminal. Reading a long message is preferable to listening to it, faster, one can skip to desired parts and one does not have to listen to the entire message again if one misses a vital piece of information the first time round. For one page units of text, such a system could operate for $25 a month, permitting large numbers of individuals to enter the publishing business for pleasure or profit. Improvements in technology will likely permit the offering of lengthier texts at slightly higher costs. Low cost home computers, already available commercially may permit the offering of classified ad and other search services with minimal capital investment. For distant communications, telephone, already cheap at off hours can be supplemented by packet networks optimized for the low cost delivery of data.

With the development of such systems, some information will be available free, except for telecommunications costs on the order of the cost of a local phone call. This will be provided by governments and non-profit institutions and under some circumstances by commercial enterprises. Governments currently sponsor telephone libraries which make recordings of medical information available to the public over the phone. Some information will be provided to serve as an advertising medium, as weather recordings in some areas of the country are preceeded by a sponsors advertising message. Some information will be provided by commercial firms to educate the public in dealing with them; a series of telephone recordings concerning the shipment of household effects made available by Washington area moving firms is an example.

The existence of the vanity press indicates that some authors typically, poets, novelists and crusaders for a cause, will make their work available to the public free and will even pay the mechanical costs involved. Authors of scholarly communications, as this paper poses to be,

expect no compensation for their intellectual labors but might balk at subsidizing distribution costs although this is a common practice through the use of page charges.

Most authors however hope to earn a living from their work and will expect ro receive payment in return for granting access to the information they have produced. Hence the need for a payment mechanism.

The absence of a mechanism to pay for information supplied over the telephone has impeded the growth of informations services using the telephone network as a delivery system. This network, a major national asset representing over fifty billion dollars has in consequence been insufficiently utilized as an information distribution medium, to the detriment of our society in general.

A trivial but illustrative example of this problem is provided by studying the availability of snow reports for skiing. For downhill skiing, in which participants must purchase a lift ticket for about ten dollars, snow reports are available from the lift operator by telephone as there is an economic incentive for their provision. Indeed, as a marketing took, recorded announcements of snow conditions are made available as local calls in the ski area's major markets to encourage inquiries. In contrast, cross country skiing requires no lift ticket or similar expenditure to participate and there is no incentive for anyone to provide snow information. As a result it is far harder to come by; one must make long distance calls to forest rangers or in many cases do without.

The value of snow reports is equal to both types of skiier, neither wants to drive several hours to discover conditions are poor and would be willing to pay for good information were there a mechanism to do so. Even the downhill skiier now getting free reports would pay for information if experience indicated that the free reports overstated the quality of snow conditions.

A payments mechanism suitable for situation such as snow reports and other information services of small importance and low value will have to have sufficiently low transaction costs to make such services economically viable.

The number of commercial information services offered and their level of usage should be a function of the minimum charge for any service down to some

absolute level where cost ceases to be a barrier to usage. If the minimum charge is ten cents, demand may exist for trivial services, say a soup receipe of the day, that will drop sharply at a minimum charge of 25 cents. As the minimum charge can be no lower than the cost of transferring funds from buyer to seller, the cost of such a funds transfer will greatly determine the scope for development of the information services industry. For the industry to develop to its fullest potential, very low cost mechanisms of funds transfer are needed.

In addition to low cost, spontaneity of use of such mechanisms is important to the full development of information services. A prospective buyer having the electronic equivalent of a checkbook or charge card should be able to obtain service instantly without prior arrangement with the vendor.

Currently, vendors of on-line services require prospective users to sign contracts with them prior to the assignment of passwords and instructions for using their system. Charges for usage are billed for subsequent payment. The contract also serves to place some restrictions on users of the information obtained from the system as desired by the system operator and the data base generator.

This system is adequate so long as the number of vendors and users is relatively small, as is the case today. However, industry growth and the technical developments noted above will bring about a proliferation of electronic publishers. The number of users, today limited to a sizable number of institutions will expand to individuals with the entire literate population a potential market. The sheer numbers of potential relationships precludes prior contracts and necessitates a system permitting spontaneous access to services.

For the maximum development of information systems, users on learning of a desired information service (this is of itself no mean problem) should be able to communicate identifying data, such as bank and account number, perhaps also a password to the device offering the service and to have instant access to the service's facilities. The charge for the service would be transferred from the users account to the sellers would appear on the users bank statement in due time.

The need for such a capability to transfer funds electronically as described above is not limited to future information services. Our existing funds transfer system, based on a paper medium, the check, is hard pressed to keep up with the rapidly increasing volume of transactions (27 billion checks in 1974) generated by our society and electronic funds transfer systems (EFTS) are seen as the solution. The development of EFTS has been the subject of considerable study by academics, lawyers, banking industry groups and the government.

As a result of the interest in EFTS by many segments of society, indications that its introduction could produce a variety of legal, economic and social impacts on our culture, and a desire to anticipate the implications of a new technology in advance of its implementation, the National Commission on Electronic Funds Transfers was established by Congress in 1974. The Commission was charged to "conduct a thorough study and investigation and recommend appropriate administrative action and legislation necessary in connection with the possible development of public or private electronic fund transfer systems"(1).

The Commission has been appointed, begun work, held hearings and produced a progress report on its work to the President and the Congress (2). It is to complete its investigation and report its findings by October 1977.

If the consumer oriented information services that will evolve in the future are to develop to their full potential, it is essential that whatever EFTS is adopted is compatible with the needs of these information services. To date however, the involvement of the information science community in the development of EFTS has been minimal. The community is not represented be membership on the Commission nor has it sought to present its needs in the Commission's hearings (3). This is unfortunate since it is to information science professionals that society must look for the vision necessary to insure that the systems adopted today will meet the needs of the information industries of tomorrow.

ASIS should consider setting up a group to monitor developments in the EFTS field, define the needs of future information services in respect to EFTS and participate in the system design process to insure that those needs are accomodated. Only thus can we insure that the information systems of the future will develop to their fullest potential in the service of our society.

(1) Public Law 93-495

(2) Programs, Plans and Accomplishments of the National Commission on Electronic Fund Transfers; National Commission on Electronic Fund Transfers, NTIS PB-258 611

(3) Witness lists, NCEFT

Cost Effectiveness of Installation of an On-Line Editing System at Chemical Abstracts Service

W. G. Cox, F. K. Salinshick, and R. L. Wigington
Chemical Abstracts Service
Columbus, Ohio

In the fall of 1975, CAS began a production experiment to evaluate the feasibility of performing all of its necessary technical content edits by way of on-line computer terminals. By year's end 100% of CA new substance Index Names were being edited on-line instead of by traditional marked copy, clerical keyboarded batch processing. The success of this effort reinforced the decision to convert the whole range of abstract/index entry editing to on-line processing. By December, 1976, 75% of all professional editorial review was being done on-line.

The economics of total on-line editing of substance names, abstracts, and index entries at CAS can be summarized as follows:

- An over-all increase in editor productivity of approximately 20%.
- Reduction of editorial support staff (keyboarders, proofists, data control, etc.) by 40 full-time equivalents, accomplished through reassignment and attrition.
- Elimination of 7 million pieces of paper annually (data sheets used for hard copy edit and correction processing).
- Currency improvement for CA of 4-5 days.
- No significant change in the amount of 370/168 computer time required to support editorial review.

The only additional costs identified to date are those associated with the acquisition and maintenance of three minicomputer-terminal clusters. These costs are small when compared to the savings in human resources.

HAS THE END USER BEEN FORGOTTEN? AN INFORMATION PROFILE OF ACTIVE RESEARCH BIOSCIENTISTS.

Dade T. Curtis
Milner Library
Illinois State University
Normal, Illinois

To learn what impact machine readable bibliographic data bases have had on the information habits of bench bioscientists, a survey of the 98 registrants at a 1976 invitational summer research conference in New Hampshire was conducted. These scientists form a suitable prototype of the active biosciences research community as conference participants are selected for research credibility or potential from an applicant pool of broad geographic, institutional, and personal character.

Researchers in the biosciences are appropriate targets for evaluative studies in market penetration of interactive search systems. Information retrieval in the life sciences has been heavily promoted via MEDLARS and MEDLINE for more than five years. Bioscientists also use computers in research (1), have a tradition of information literacy (2-5), and tend to be individualists in their information handling (6). As Wanger, Cuadra, and Fishburn (7) categorize only 19.2% of all researchers as "not served" by on-line search systems, bioscientists could be expected to be among the heavier users.

DATA COLLECTION

A four page questionnaire was developed, tested, and mailed with an explanatory cover letter to every person on the conference registration list. Oral interviews with nine registrants had previously shown the feasibility of this survey.

Responses, obtained from 60% of the registrants, came from 18 states, 5 foreign countries, 20 medical schools, 7 graduate universities, 3 non-profit research institutes, 2 government laboratories, and 1 industrial research center. The total registrant population was compared with the respondent sample to determine that a valid return was obtained. No significant variance was found when geographic origin, institutional affiliation, age of scientists, or areas of research activity were compared.

CHARACTERISTICS OF RESPONDENTS

A relatively youthful, well educated group of 57 men and 1 woman makes up the sample. All but two persons have doctoral degrees; these two are candidates for doctorates. While 79% are less that 46 years of age, the mean age is 38 years and the median age is 35 years. Sample age level correlates with the finding that 65% of the respondents decided to pursue their scientific careers in the past 15 years. It also explains why

43% of the respondents characterize themselves as "emergent" scientists rather than as "established" scientists.

Respondents are productive scientists and are esteemed by their peers: 86% have published more than once in the past six years and 82% have obtained one or more grants since 1970. Students, postdoctoral fellows, and persons sited outside academe make up the unpublished-unfunded minorities. Research fields, in decreasing order of number of replies, are physiology, biophysics, biology, pharmacology, physics, cardiovascular medicine, mathematics, biochemistry, and bioengineering.

Computer technology is used by 84% in their research activities and by 55% in their teaching or administrative duties. Yet 57% of the respondents have *never* searched a computer file, either on-line or batch. This is remarkable when you consider that 51% are members of medical faculties and an additional 35% are members of other university faculties where search services are available.

INFORMATION PATTERNS OF RESPONDENTS

Sources of Information. Very few of the respondents search to fill either their current or retrospective information needs (Fig. 1). Interactive searching is a primary information source for only one person's current needs and a primary information tool for the retrospective needs of only five respondents. Most respondents obtain their information from either conventional formal sources or through informal communication processes.

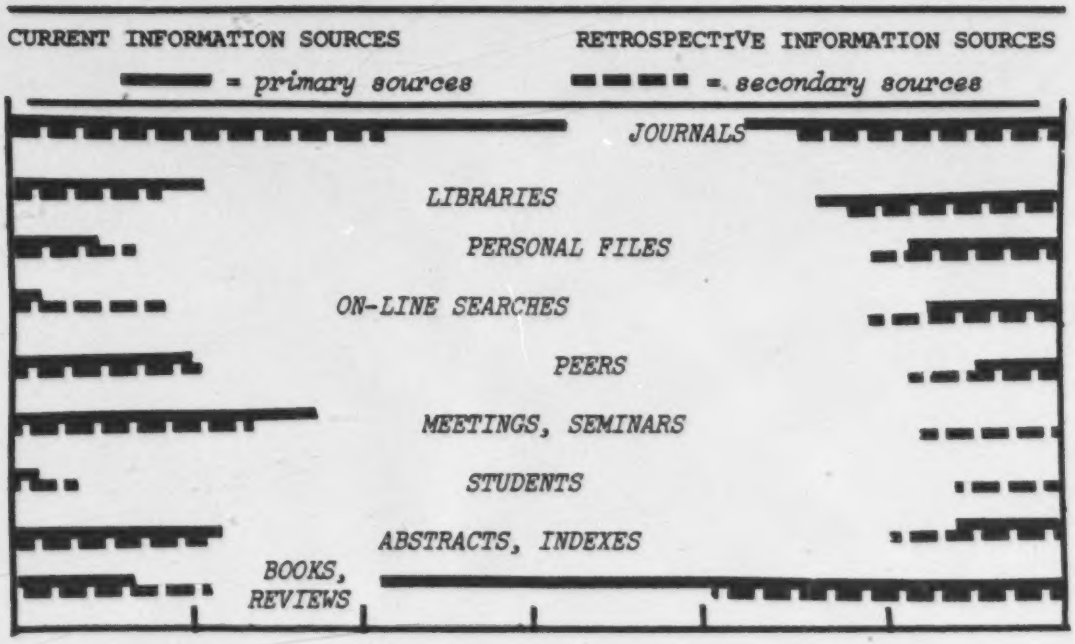


Figure 1. Sources of current (left) and retrospective (right) information for survey respondents. Divisions along the X-axis represent 10% of total replies for each category. Origins are found at lower left and lower right corners.

Because Fig. 1 implies that few of the respondents in this survey are information "gatekeepers" in the greater biosciences world, other explanations to explain their passive searching behavior must be found. Little or no correlation is present in this group of scientists between non-searching and between knowledge of the search process, availability of search centers, areas of research, age of individuals, and money needed to pay for searches.

User Competence. Non-searching cannot be tied to an overwhelming satisfaction with the individual efforts of the respondents. When asked to evaluate their success in obtaining information, a simple majority feel proficient only in research (Table 1). In other activities a general dissatisfaction with their personal attempts is seen.

Activity	Expressed Level of Success, % of sample.			Totals
	Good	Adequate	Poor	
Research	57.1	34.7	8.1	99.9
Learning	38.8	32.7	28.5	100.0
Non-professional interests	30.6	32.5	36.9	100.0
Teaching	28.6	38.8	32.6	100.0
Administration	8.2	28.6	63.1	99.9

Table 1. Respondents' subjective evaluation of their personal success in obtaining information.

Comparison of Searchers, Potential Searchers, and Non-Searchers. Searchers are all "emergent" American scientists. None searches extensively. While no data exist here to support the presumption, it is possible that these people are not yet full members of their respective "invisible colleges" and hence are receptive to time saving alternates in information retrieval. Potential searchers come from varying ages, ranks, and nationalities. Some have searched but many have not. Some 88% complain either of the coverage, complexity, indexing policies, diversity, or lack of personal specificity in the printed files and machine systems with which they are familiar. More than 63% demand personal control over their information processes. Again we learn the wisdom of consulting the user before systems are designed.

The single common factor between searchers and potential searchers

is a general understanding of the publishing and/or information industries.

Non-searchers have an identity commensurate to the potential searchers. A poor understanding of current publishing, information, and political trends tends to characterize them. The group operates within the status quo and sticks with information habits learned in graduate school. Changes desired by this group include fewer journals, fewer symposial volumes, better libraries, codified reviews, and easier contact with their colleagues.

MARKETING IMPLICATIONS

These data strongly suggest that past and current strategies used to market on-line search services to active research bioscientists employed largely in academic environments have not been markedly effective. It will be contended that this type of end user has not been forgotten, but suffers from benign neglect.

A number of interrelated elements may be responsible. First, an intriguing market paradox exists in that we buy, sell, and supply information to the very persons who initially produce and eventually consume the product. Second, communication barriers, always a weak link, are likely when producers and consumers interact with middlemen. Third, as few of our intermediaries are adequately trained both in modern marketing techniques and in the basic or derived product, even fewer of the producer-consumers have been convinced of the pragmatic utility of searching. Fourth, strong competition from other methods of demonstrated efficacy fractionates the user clientèle.

Each of these elements can be argued until the cows come home, visit, and leave. To avoid an agricultural confrontation, several one liners will instead be suggested as possible actions to remedy the problems:

1. Analyze your market. Offer files which fit contemporary use and value patterns.

2. Publicize the availability of reviews and books in existing data bases. Better, form a merged file of reviews and books from several complementary data bases.

3. In the absence of a single "American Biological Society," exhibit at and cooperate in the meetings of small to medium sized societies.

4. Woo society officers and journal editors for time and space to present searching philosophy and examples.

5. Use your contacts in university towns.

6. Use media in university towns.

7. Produce user operable, self-learning, simple instructional packages.

8. Capitalize upon any and all inquiries.

9. Concentrate on emergent bioscientists.

10. Leave your terminal or mainframe and mingle with your customers.

11. Educate and train your advance salemen and saleswomen, many of whom are librarians, in basic marketing logic and in productive selling methodology.

12. Encourage and support cooperative teaching ventures by information intermediaries and science faculty members.

If we, as an information industry, have not yet recognized that these are fundamental tactics, we should.

REFERENCES

1. 1974. "Conference on the Computer as a Research Tool in the Life Sciences." Papers and remarks from the 1974 Annual Meeting of the Federation of American Societies for Experimental Biology, Aspen, Colorado. *Federation Proceedings*. 1974; 33(12): 2311-2421.

2. Badre, A. L. et al. 1974. "Science Information Transfer for Learning." *in* Zunde, P., Ed., *Information Utilities*, Proceedings of the 37th ASIS Annual Meeting, Atlanta, Georgia, 13-17 October 1974. Washington, D.C.: American Society for Information Science. 1974; 120-124.

3. Council on Biological Sciences Information. 1970. *Information Handling in the Life Sciences*. Washington, D.C.: Division of Biology and Agriculture, National Research Council. 1970. 79 p. ED 040 722.

4. 1974. "Biomedical Literature and Information Services." Eleven papers prepared as a Special Topic. *Federation Proceedings*. 1974; 33(6): 1693-1723.

5. Steere, W. C. 1976. *Biological Abstracts/BIOSIS; the First Fifty Years. The Evolution of a Major Science Information Service*. New York: Plenum. 1976. 233 p.

6. Blaxter, K. L.; Blaxter, M. L. 1973. "The Individual and the Information Problem." *Nature*. 1973; 246(5432): 335-339.

7. Wanger, J.; Cuadra, C. A.; Fishburn, M. 1976. *Impact of On-Line Retrieval Services: A Survey of Users, 1974-75*. Santa Monica, Calif.: System Development Corporation. 1976; p. 36.

THE VALUE OF SCIENTIFIC AND TECHNICAL INFORMATION TO DEVELOPING AREAS AND PROBLEMS IN TECHNOLOGY TRANSFER

Richard A.V Diener

Xerox Corporation
Webster, NY 14580

ABSTRACT

Scientific and technical information is a subset of the total information contained in a society; a society that is interpreted as an informational phenomenon. The "advanced-ness" or "modern-ness" of a society is reflected in the specificity and complexity of its most advanced concepts. Value is a human decision based upon socio-cultural and personal needs and expectations. In Western society, value is often equated with cost (or economics) while in non-western societies there are more parameters to value (Aristotle suggested seven: aesthetics, economics, ethics, morality, politics, sociality, spirituality). Transferred information must impact on the existent information systems in such a way as to blend with the indigenous concepts, their relationships and values. One view of the appropriateness of information is the interaction of *concept specificity*, *concept clustering*, *concept relationship* and *concept use analysis*. Technology transfer activities are considered as information dissemination between vastly different socio-informational systems. The problems encountered by foreign aid personnel in having the technologies accepted by recipient groups is viewed as a problem in the transfer of inappropriate types of information. The framework of a methodology for transferring information to less developed users is discussed.

> [A]ll observers are not led by the same physical evidence to the same picture of the universe, unless their linguistics [informational[1]] backgrounds are similar, or can in some way be calibrated.
>
> Benjamin Lee Whorf, 1956
> *Language, Thought and Reality*

INTRODUCTION

All to frequently the value, efficiency, effectiveness and benefit of information are mirrored against *cost* only. But considering the overtly economic nature of western society, this is an understandable gauge. In foreign aid programs

> aid includes direct government-to-government aid, and also aid given through international organizations..., foreign private investment..., private charities ... etc. and are viewed as ... the flow of public *finance* to *poor* countries. [*italics mine*] [1, p. 13]

However, the rural recipients of technology transfer activities in developing, often third world, nations is still concerned with value that can not be measured in dollars or satisfied by money.

Value was conceived of by Aristotle[2] as composed of relatively equal portions of aethetics, ethics, morality, politics, sociality, spirituality and economics. Any study of value which negates or overtly surpresses any of these components would be an injudicious effort. Those transferring technology to developing nations place high value on their underlying goal, the maximal up-grading (progress) of the recipient's technical "know-how"[3] in a minimum of time. As an end result of this goal, these activities often attempt to transfer inappropriately sophisticated technologies which cause massive alterations in the local social structure over very short periods of time; short relative to normal social change and cultural evolutionary processes. These technologies often fail to support the expected continuation of their rapid growth because of a lack of peripheral, supporting information in the culture. This discussion will deal with the interaction of technology and society as socio-informational phenomena, and particularly with the effects of the direct infusion of technological information into underdeveloped informational settings.

TECHNOLOGY AND SOCIETY AS INFORMATIONAL PHENOMENA

Culture, an anthropological term, has been defined previously as the information content of a social unit which is transmitted intergenerationally by learning [2,3]. Social information has the additional features of being distributed statistically amongst the members of that social unit and subject to alterations in content and structure through time. The ultimate significance of information in a social setting is its usefulness in problem-solving and decision-making. These activities can range over the very broad spectrum of sophistication from as simple an operation as knowing how to use an ignition key to start a tractor to as complex an operation as designing, building and remotely operating an interplanetary spacecraft. The relatively of this sophistication is socio-culturally founded. Any society, whose average member has difficulty in maintaining a tractor, is highly unlikely to put a spacecraft on Mars or to be able to service and fly F-114 fighter planes on their own. The difficulties that have been encountered in training people

to supply technical service support for transferred military technology to underdeveloped nations like South Vietnam and Iran have been adequately reported by the television news media.

At its ultimate reduction, technology is an informational phenomenon, containing information about form, dimensionality, function, operation, etc.; and is a subset of the total socio-informational picture that is society. Most technology exhibits the limiting criterion of being physical, a tool or machine; in anthropological parlance, a material culture. Items of material culture are operationalized in a socio-informational setting by knowing how to utilize, repair, design, build, modify and maintain them. All of these activities require information of various sorts. Lauriston Sharp (4) discussed the knowledge required to prepare stone axes by the Yir Yoront, a very primitive Australian aboriginal group.

> The production of a stone axe required a number of simple skills. With the idea of the axe in its various details well in mind, the adult men -- and only the adult men -- could set about producing it, a task not considered appropriate for women and children. First of all, a man had to know the location and properties of several natural resources found in his immediate environment: pliable wood, which could be doubled or bent over the axe head and bound tightly to form a handle; bark, which could be rolled into cord for the binding; and gum, with which the stone head could be firmly fixed in the shaft. These materials had to be correctly gathered, stored, prepared, cut to size, and applied or manipulated. They were plentifully supplied by nature, and could be taken by a man from anyone's property without permission. Postponing consideration of the stone head of the axe, we see that a simple knowledge of nature and of the technological skills involved, together with the possession of fire (for heating the gum) and a few simple cutting tools, which might be nothing more than the sharp shells of plentiful bevalves, all of which were available to everyone, were sufficient to enable any normal man to make a stone axe. [4, p. 129-130]

SOPHISTICATION AND APPROPRIATE TECHNOLOGY

One of the principle parameters of information from a social perspective is its sophistication; interpreted informationally by *concept specificity* and *concept relationship analyses*. The more sophisticated or advanced a society is, the more specific are its most advanced concepts. For example, the most advanced conceptualization of a tractor for a rural farmer in an underdeveloped area might be its function as a tool for plowing the ground while his counterpart in an advanced society might well be an accomplished tractor mechanic with a full understanding of all the concepts involved in the engine, transmission, drive train, electrical system, accessories, etc., as well as a functional awareness of the operational interrelatedness of all these subsystems. This lack of understanding on the part of the rural farmer severely, and sometimes critically, limits his ability to utilize the technology. What is probably worse yet is that this lack of supporting *technical* information not only impacts on the *economic* value of the machine, that is directly inhibits the use of the machine, but also impacts on the *other social* values as well.

The other six (aesthetic, ethical, moral, political, social and spiritual)

values frequently suffer from a lack of supporting information in other than the technical realms of the total socio-informational setting. Advanced technologies are nested in a whole spectrum of social expectations, behavioral assumptions, cultural traits, common knowledge, much of which is totally absent from the cultures of developing nations. In bureaucratic terms, it suffers from a lack of infrastructure such as reliable utilities, materials, energy, transportation systems, etc. This lack of cultural predispositions, expectations and assumptions further inhibit the total usefulness of this transferred technology.

A severe example of the problems which can be caused by the inappropriate transfer of technology (a sibling concept to the transfer of inappropriate technology) was discussed in 1952 by Lauriston Sharp (5). As stated above, the knowledge needed to make, therefore to own, stone axes was restricted by custom to the males of the Yir Yoront. As a gesture of friendship in making contact with them, Christian missionaries distributed steel axes. But due to their poor understanding of the Yir Yoront's group structure and their customs about this type of technology, the missionaries gave axes to males, females and children alike. As the axe was virtually the only technology of this group and was used through the restriction of information as a social status control mechanism by age and sex, this indiscriminant distribution caused an evaporation of much of the social structure.

> For the Yir Yoront ... personal deprivation and frustration in a confused culture would produce an overload of anxiety. The mythical past of the totem ancestors would disappear as a guarantee of a present of which the future was supposed to be a stable continuation. Without the past, the present could be meaningless and the future unstructured and uncertain. Insecurities would be inevitable. [5, p. 22] ... changes in overt behavior, in technology and conduct, weakened the values inherent in a reliance on nature, in androcentrism or the prestige of masculinity, in age prestige, and in the various kinship relations. A scene was set in which a wife or young son, his initiation perhaps not even yet completed, need no longer bow to the husband or father, who was left confused and insecure as he asked to borrow a steel axe from them. For the woman and boy the steel axe helped establish a new degree of freedom which was accepted readily as an escape from the unconscious stress of the old patterns, but which left them also confused and insecure. Ownership became less well defined, so that stealing and trespass were introduced into technology and conduct. Some of the excitement surrounding the great ceremonies evaporated, so that the only fiestas the people had became less festive, less interesting. Indeed, life itself became less interesting, although this did not lead the Yir Yoront to invent suicide, a concept foreign to them. [4, p 135]

Peter Berger (6) echoes the same sentiments when he says "... the loss

> The loss of collective and individual security carries with it the constant threat of isolation as well as meaninglessness. In their cumulative result, these processes add up to a pervasive condition of "homelessness" -- man is no longer "at home" in society, in the cosmos, or ultimately with himself. [6, p. 34]

The Yir Yoront are an hyperbolic example of the type of disruption

that frequently occurs when societies of vastly different informational size, content and sophistication are brought into contact and information begins to flow. Information, behaving like a fluid under pressure, streams from the more to the less sophisticated society, often inundating more primitive socio-informational structures and disrupting even the more sophisticated societies though less severely. As a result, man's advanced technologies and information endanger and extinct less sophisticated societies with the same abandon and thoughtlessness that his industrial pollutants endanger and extinct hapless species that stand in the way of progress and profit.

Many words have been written and anxious tears have been shed as a result of *our* recent information explosion and its subsequent information overload (and we were as prepared for it as any society could have been). And yet we distribute the technologies which have resulted from this explosion to hapless and unsuspecting people all over the underdeveloped world in the name of peace as aid. Considering the rates of mental illness, delinquency and crime in western societies, conditions which can be attributed as least in part to this uncontrolled information explosion and its resultantly produced alienization and anxiety amongst our people, one shudders at the thought of its potential, and unfortunately in some cases real, effect on the unsophisticated of the third world. In talking about magic and religion in New Guinea, Philip Newman (7) stated that ...

> All men experience anxiety when confronted with situations they cannot control, and in many societies natural methods of influencing and predicting events work only part of the time. In such instances, supernatural forces are invoked to account for such events and man's relation to them.... the use of magic and witchcraft by a highland New Guinea people ... shows that they employ such practices thoughtout their lives whenever faced with uncertainty.... that magical procedures can be ranked according to their ability to release tension, and that the choice of particular magical practices correlates with the degree of anxiety to be reduced. [7, p. 316]

It is our responsibility -- our ultimate responsibility for our own well being as a future lesser developed country -- to insure the sanity of the direction these third world countries will take in their maturation processes. As the decline of power in western and communist countries is inevitable, and in fact can already be seen, many of the world leaders of the twenty-first century will come from the third world. The legacy we leave them will be spun in the information we disseminate to them and the ways in which it is presented.

There is, potentially as least, a solution to the exceedingly dismal prophecy of continued human degradation on an international scale. It lies the selective transfer of appropriate technology or the selective dissemination of appropriate information.[4] Appropriateness is closely allied with value, total value. As it is appropriate only if it is valuable, its value is measured by the degree to which it is supportive of the extant population. For a technology to be appropriate it must do more than produce higher output than whatever it is degigned to replace, it must be totally comprehensible to the people receiving it; comprehenible in form, function and, of course, repairability. It must fit into the total social milieu and be sophisticated enought to better the life of the people receiving it while still being emotionally, aesthetically, ethically, morally, socially, politically and

spiritually satisfying. If it fails to be satisfying, it is worse than useless as it will produce anxiety and prove disruptive to the recipient society.

TRANSFERRING APPROPRIATE TECHNOLOGY

The British economist, E.F. Schumacher, along with his Intermediate Technology Development Group, Ltd. and other agencies, have worked for a number of years to transfer appropriate technology to developing areas of the third world, most notably for Schumacher and associates to rural India, with very encouraging results (8). Yet they seem to be working from a semi-intuitive, seat-of-the-pants approach to designing the intermediate technology. It is felt that a socio-informational approach to the analysis of the problem of identifying the "appropriateness" would prove fruitful. By performing analyses of *concept specificity, concept clustering, concept relationship* and *concept use* both of the potential recipient group in real time ethnographically and of the technology proposed for transfer in a historical or developmental framework, the organization transferring the technology could match these levels of specificity, interrelatedness, clustering and use to provide a technology which is acceptable to the recipient group, as well as beneficial to their needs; therefore of maximum value.

The methodology that would be used to perform the analyses of *concept specificity* and *concept clustering* have their theoretical groundings in classification research. Most classificatory schemes have the relation of hierarchy imbedded within them. The theoretical groundings for *concept relationship* would be found in the work of Jason Farradane on relational analysis (9), and possibly in S.R. Ranganathan's Colon system. There are six categories of *concept use* which cover the basic interrogatives of what?, when?, where?, who?, how? and why?. These recognize that information needs to be understood as a problem-solving and decision-making entity which responds to these types of use patterns.

CONCLUSION

The value of scientific and technical information lies in its usefulness and acceptance which in turn rely on its appropriateness. The parameters which influence *value* are deeply imbedded in the socio-cultural structure, and exceed simply an economic interpretation. Information pervades the totality of society like an intricately woven fabric; tear at the fabric and pull at the threads, and it will fray and tatter.

BIBLIOGRAPHY

1. Little, I.M.D.; Clifford, J.M. 1966. *International Aid.* Chicago: Aldine; 1966.
2. Diener, R.A.V 1976. "Information systems in human cultural evolution". *American Society for Information Science.* 5th Annual Mid-Year Meeting, Nashville, TN, 1976, Proceedings, Volume 5; 209-218.
3. Diener, R.A.V 1976. "An information-based hypothesis for the evolution of socio-political power". *American Society for Information Science.* 39th Annual Meeting, San Francisco, CA, 1976, Proceedings, Volume 13; 75.

4. Sharp, Lauriston. 1971. "Technology and social institutions: Australia". In: *Exploring the Ways of Mankind*, 2nd ed., Walter Goldschmidt, ed. New York: Holt, Rinehart and Winston; 1971. 128-135.

5. Sharp, Lauriston. 1952. "Steel axes for stone-aged Australians". *Human Organization* 11:17-22.

6. Berger, Peter L. 1974. *Pyramid of Sacrifice: Political Ethics and Social Change*. New York: Doubleday; 1974.

7. Newman, Philip L. 1974. "When technology fails: magic and religion in New Guinea". In: *Conformity and Conflict*, 2nd ed.; James P. Spradley and David M. McCurdy, eds., Boston: Little-Brown; 1974. 316-325.

8. Schumacher, E.F. 1973. *Small is Beautiful: Economics as is People Mattered*. New York: Harper-Row; 1973.

9. Farradane, Jason E.L. 1950. "A scientific theory of classification and indexing and its practical applications". *Journal of Documentation* 6:83-99; and 1952. "A scientific theory for classification and indexing: further considerations". *Journal of Documentation* 8: 73-92.

NOTES

1. Language is the method for communicating and expressing information and knowledge.

2. Aristotle's categories are chosen not to reflect the value structure of all or any developing nation but only as a model of value structures. In any specific instance where the value structure of the society in question is known, it should be used in the value analysis.

3. "Know-how" is different from "know-why"; the former deals with operation and functioning while the latter involves meaning and understanding. Both are necessary for a full understanding of a technology or any information.

4. Admittedly, this is only a partial solution. The application of the basic method to developing bureaucratic structures would further assist in the development/maturation process.

VALUE: DATA, INFORMATION, KNOWLEDGE

Richard A.V Diener
Xerox Corporation
Webster, NY 14580

INTRODUCTION

The phenomenon generically known as *information* can be re-interpreted into three phenomena: *data*, *information* and *knowledge*, when viewed through Concept Use Analysis. Data are sensory and perceptual phenomena, while information and knowledge are conceptual phenomena interpreted as structural and functional, respectively. Concept Use Analysis is a method for identifying information and knowledge by the categories of questions they can answer. It is assumed in this discussion that the final value of information is in its usefulness in answering questions.

CONCEPT USE ANALYSIS INTERPRETATION

Concept Use Analysis recognizes six basic questions which can be asked and subdivides them into structural questions (What?, When?, Where? and Who?) and functional questions (How? and Why?). The structural questions are answered by *information* while functional questions require *knowledge*. As both *information* and *knowledge* are conceptual phenomena, they are then considered cognitive level activities and *data*, being a sensory-perceptual phenomenon, is considered a sensory level activity.

Data is considered a category of "information" which deals with the sensory impressions received by the senses -- visual, auditory, tactual, taste and olfactory. These impressions are encoded neurochemically as a perception which is retrievable upon demand (when recollected). These perceptions supply the raw material for the cognitive processes (thinking). The interpretations of these perceptions, their conceptualizations, are socio-culturally defined phenomena, shared by the group experiencing the data. Once sensed, encoded, perceived and involved in a conceptualization, data has been altered into a new form ... information or knowledge.

Information is that group of conceptualized data that deals with the structure of a phenomenon; particularly as it relates to being, temporal and spatial-dimensional components, and personalities. These are roughly equivalent to the interrogative categories: What?, When?, Where? and Who?. It might be said that Information identifies the psycho-physical components of the universe, those which deal with time, space, identity and personality.

Knowledge is that group of conceptualized data that deals with the function and meaning of a phenomenon; including its operational parameters. These are roughly equivalent to the interrogative categories: How? and Why?. It might be said that Knowledge identifies the energy/cognitive components of the universe. Knowledge tends to be more complex than information, requiring more background understanding.

VALUE OF DATA, INFORMATION AND KNOWLEDGE

Value lies in the eyes of the beholder. Only the user of an information system/service/product can assign a value to the output from that system/service or of that product. One of the major components of a user's decision about value is the user's satisfaction with that system/service/product. One source of this dis-satisfaction, and therefore low perceived value by the user, is the failure of the system/service/product to provide the proper answer to the questions presented by the user. To overcome some of this dis-satisfaction, a more effective information system design is proposed. In it, the type of question/need, interpreted by an appropriate interrogative pronoun or adverb, is matched to information or knowledge in a data base that is encoded for retrieval using interrogative categories.

It is the mission of this table to investigate the value of data, information and knowledge. The above is given as a stimulus.

NUMERIC DATA BASE SYSTEMS : CONCEPTS AND CAPABILITIES

Wayne D. Dominick
University of Southwestern Louisiana
Computer Science Department
P. O. Box 4-4330
Lafayette, Louisiana 70504

ABSTRACT

As practitioners in a research-intensive discipline, the members of the scientific community have not only witnessed, but also participated in the constantly expanding information explosion. In response to the need for storing, retrieving and analyzing large masses of interrelated numeric and textual data, increasing emphasis is being placed on the utilization of generalized data management systems to support integrated data bases.

In order to lay the groundwork for the discussion of numeric data bases in specific scientific application areas, this paper will provide an overview of the concepts and capabilities of numeric data base systems. The presentation of the material is intended as a tutorial survey of the features of such systems rather than as a detailed examination of any one specific system or specific application area.

Topics to be addressed include data base design and data base definition, focusing on specifying data item names, types, units, validation criteria, groupings and interrelationships; data base interrogation; computational processing, including statistical data analysis; report generation, including graphical display; and the dynamic aspects of data base redefinition.

INTRODUCTION

The design of numeric data base systems is predicated upon the desire to support several different classes or categories of data base system users. Such classes include data base administrators, application programmers, non-programming end users and parametric users.

Data base administrators are the individuals who maintain centralized control and managerial responsibility for particular data bases. Data base administrator responsibilities include determining the desired information content of a data base, defining the data items and data interrelationships to be represented within the data base, specifying the data security rules and data access authorizations to be employed, maintaining liason with the users of the data base, and redefining or restructuring the data base as needs demand, either for optimized

NUMERIC DATA BASE SYSTEMS : CONCEPTS AND CAPABILITIES

Wayne D. Dominick
University of Southwestern Louisiana
Computer Science Department
P. O. Box 4-4330
Lafayette, Louisiana 70504

ABSTRACT

As practitioners in a research-intensive discipline, the members of the scientific community have not only witnessed, but also participated in the constantly expanding information explosion. In response to the need for storing, retrieving and analyzing large masses of interrelated numeric and textual data, increasing emphasis is being placed on the utilization of generalized data management systems to support integrated data bases.

In order to lay the groundwork for the discussion of numeric data bases in specific scientific application areas, this paper will provide an overview of the concepts and capabilities of numeric data base systems. The presentation of the material is intended as a tutorial survey of the features of such systems rather than as a detailed examination of any one specific system or specific application area.

Topics to be addressed include data base design and data base definition, focusing on specifying data item names, types, units, validation criteria, groupings and interrelationships; data base interrogation; computational processing, including statistical data analysis; report generation, including graphical display; and the dynamic aspects of data base redefinition.

INTRODUCTION

The design of numeric data base systems is predicated upon the desire to support several different classes or categories of data base system users. Such classes include data base administrators, application programmers, non-programming end users and parametric users.

Data base administrators are the individuals who maintain centralized control and managerial responsibility for particular data bases. Data base administrator responsibilities include determining the desired information content of a data base, defining the data items and data interrelationships to be represented within the data base, specifying the data security rules and data access authorizations to be employed, maintaining liason with the users of the data base, and redefining or restructuring the data base as needs demand, either for optimized

execution efficiency or in response to new requirements of the user community.

Application programmers represent programming support personnel whose function is to program extensions to the data base system to provide application tailored processing not provided within tne data base system. They represent a class of users with a significant amount of programming expertise. In contrast, non-programming end .users represent information specialists, concerned witn the semantics and information content of the data rather than with the internal functioning of the system and programmed extensions to the system. Finally, parametric users represent a class of users whose primary function is to invoke pre-defined procedures over a data base in a routine manner. Such users are often support staff, being neither information specialists nor system specialists.

In order to support the needs of these different classes of users, numerous levels of data base system capabilities are required. The primary intent of this paper is to provide a general overview of these capabilities, focusing on the support required for managing and accessing numeric data bases.

DATA BASE DEFINITION

Data base definition is the process of determining tne desired information content of the data base and defining tne structures and attributes of all entities to be stored witnin the data base. Data base definition is tne function of the data base administrator.

Data base definitional capabilities required include facilities for defining the following types of data base characteristics:

- Data names
- Data types
- Data value synonyms
- Data value units equations
- Data validation criteria
- Data access locks
- Data input/output conversion/editing specifications
- Data groupings
- Data ordering relationships
- Data indexing

Data names provide the facility whereby each individual data item witnin a data base can be referenced within a data access request. Data types define the general value content of data items as numeric-valued (integer, decimal), character-valued (alpnanumeric, alphabetic, non-numeric), date-valued, coded and so on. Data value synonyms define the allowable synonyms for specific data values and data value units equations define the units in which data values are stored and the transformations to be used in converting from one unit of measure to another, e.g.,

from inches to feet to miles to kilometers to meters to centimeters. Data validation criteria define the allowable values for a particular data item, either as a range of data values, a set of data values or via a user-supplied data validation procedure for more complex data validation requirements.

Data access locks define data item security, i.e., who is allowed to retrieve the data item, update the item, delete the item, add new values for the item, etc. Data access locks may consist of password specifications, user name or identification number access control lists or more sophisticated data base administrator defined access validation procedures. Data input/output conversion/editing specifications define the transformations to be performed on the data values as they are input to the system or as they are output from the system. Data groupings define the aggregation of data items into groups and data ordering relationships define the ordering of instances of data items within such groups. Finally, data indexing defines the types of indices into the data base to be generated from the values of specific data items.

INTERROGATION SPECIFICATION

Interrogation is the process of selecting and extracting data satisfying specified criteria. The basic selection construct employed is the relational expression of the following general form:

<operand> <relational operator> <operand>

where an operand may be a data item name or a data value (numeric value, character string or computed variable) and the relational operator may be any one of the standard operators LT, LE, EQ, NE, GE or GT, consistent in type with the specified operands.

Numerous variations of the basic selection construct exist. String matching options may be provided for exact match, sub-string match, root match or utilization of a universal match character. Existence checking (a data item value exists or does not exist) may be provided and compound conditions, as logical combinations of simple conditions, may be specified using the standard logical (Boolean) operators AND, OR and NOT.

COMPUTATIONAL PROCESSING

Several levels of computational capabilities exist within numeric data base systems. These include basic arithmetic computation, standard numeric functions, user-defined numeric functions and computational program interfaces.

The most basic level of computational processing involves the use of the arithmetic operations addition, subtraction,

multiplication, division and exponentiation. A second level of computational processing corresponds to the availability of standard or built-in functions. The specific functions provided by default within numeric data base systems vary widely, however, most numeric data base systems provide standard functions such as COUNT, MINIMUM, MAXIMUM, SUM, MEAN, STANDARD DEVIATION and CORRELATION. A third level of computational capability is the extension of the concept of standard functions to support user-defined functions. Via a macro definition language (procedure definition language, profile definition language) available within the data base system, individual users can define their own functions for performing tailored computational processing, save these functions and execute them whenever necessary.

The above three levels of computational processing correspond to support provided within the data base system itself. The fourth level to be discussed corresponds to computational processing performed external to the data base system. In particular, this level represents the availability of generalized linkages or interfaces between the data base system and external computational programs or systems. For example, a user may retrieve data, store these data in a generated data set and invoke some external program, such as a statistical package, for performing detailed statistical analyses on the retrieved data. The existence of such interfaces significantly enhances the processing power available to the data base system user. Not only can such a user employ all of the facilities programmed into the data base system, but also he can, when necessary, invoke any number of external programs to perform processing not provided within the data base system.

REPORT GENERATION

For reasons of flexibility, analogous to those discussed within the previous section on computational processing, numeric data base systems invariably provide their users with several levels of report formatting options. The most basic level of report generation is the use of the system default report format, wherein report content and formatting is pre-defined by the system. The logical extension of system default report formatting is user-defined report formatting. Utilizing a report definition language, a user can exercise complete control over report specifications, including:

- Report content: specifying the data items, character strings and computed variables to be included within the report.
- Positional formatting: specifying the tabbing, spacing, alignment, headings, footings, page width, page length, pagination and so on to be used within the report.
- Summarization statistics: specifying the inclusion of summarization statistics such as vertical and horizontal totals and subtotals, minimum, mean, and maximum values and so on.

- Sorting: specifying the order of entries in the report by sorting on multiple keys of any data type.

Having defined such a report, the user is invariably provided with the capability to save the report definition for future use.

A final level of report generation is the availability of alternate output formats such as graphical displays for aggregation, trend and extrema analysis (histograms, plots, bar charts, pie charts, etc.), binary output files as generated data sets and computer output to microfilm.

DATA BASE MAINTENANCE AND REDEFINITION

Data base maintenance is the process of adding new values to a data base, deleting existing values in a data base or changing existing values in a data base. The sophistication of the data base maintenance provided by a data base system depends upon several factors, including:

- Update mode
- Update data format
- Update history maintenance
- Update utilities provided
- Update access control

Update mode may be unconditional, requiring the user to explicitly specify the data base record to be updated, or conditional, allowing the user to employ a conditional expression to identify the data base records to be updated. For example:

```
Unconditional Mode:  GET RECORD NUMBER 1  <update data>
Conditional  Mode:   IF ITEM1 GT 20000 THEN  <update data>
```

Obviously, conditional updating is considerably more powerful and flexible than unconditional updating. In a similar manner, update data format may be predefined, requiring the user to explicitly specify the new data value, or computed, allowing the user to employ a computational expression to define the new data value. For example:

```
Predefined Format:  CHANGE ITEM1 TO 22000
Computed   Format:  CHANGE ITEM1 TO ITEM1 * 1.1
```

Update history maintenance refers to the availability of automatically generated transaction logs or audit trails; update utilities include facilities for automatic data base backup and checkpoint/restart processing; and update access control refers to the sophistication of the facilities provided for supporting multi-user, concurrent update processing.

In contrast to data base maintenance, which involves modifications to the values within the data base, data base redefinition encompasses modifications to the definition of the data base. Redefinition may result in completely new data items

PRESENT AND POTENTIAL VALUE OF COMPUTER COMMUNICATIONS IN INFORMATION SCIENCE

Dr. Tom Featheringham
Computer Conferencing and Communications Center
New Jersey Institute of Technology
323 High Street
Newark, New Jersey 07102

Descriptors: Computer Conferencing
Computer Communication
Information System Evaluation

This paper first describes the comprehensive communication environment provided by the Electronic Information Exchange System (EIES). Within the general context of computer communication systems the EIES is unique in that it embodies a variety of communicative functions and that its first sets of users comprise small groups of research specialists (i.e., invisible colleges). The Computer Conferencing and Communication Center at New Jersey Institute of Technology built the EIES and is currently evaluating its acceptance and use by these groups. Preliminary analysis of the system's use and acceptance is described. Some of the economics of computer conferencing systems have been reported elsewhere and are summarized in the paper.

The last part of the paper reflects on several areas where computer communication systems might have greatest impact on future information system users. Among these are the potential automation of the authorship, peer review, and referencing activities within the normal article publishing process. Also, the coupling of computer conferencing systems with on-line storage and retrieval systems will provide the future information seeker manifold knowledge resources.

Bibliography:

Day, Larry, "Computerized Conferencing: An Overview", Views from the ICCC 1974 (ed) N. Macon, International Council for Computer Communication, P. O. Box 9745, Washington, D. C.

Turoff, Murray, "The Cost and Revenues of Computerized Conferencing", Proceedings of the 1976 ICCC, Toronto, Canada.

being defined for the data base or changes to the units equations, ordering relationships, access locks, validation criteria or structural representation of existing items or groups of items within the data base. Such redefinitions may be required for several reasons. The data base needs of the user community may have changed, thus necessitating the storage of additional items of data; the natural growth of data may have caused a deterioration in time required to search the data base, thus necessitating a restructuring of the data base or the addition of certain indices into the data base; or some logical errors in the original data base definition may have been uncovered, thus necessitating their correction. In dynamic application areas, such as areas involving exploratory research where data storage and manipulation needs are constantly changing, facilities for data base redefinition assume a critical role.

SUMMARY

This paper has presented a brief overview of the concepts and capabilities of numeric data base systems. The intent of this effort has been to survey the types of features found in many existing numeric data base systems rather than to focus on any one such system.

In order to put the topics discussed within this paper into perspective with respect to generalized data base management systems and information storage and retrieval systems, the interested reader is refered to two excellent, detailed feature analysis reports: one treating the features of generalized data base management systems [1]; the other treating the features of interactive retrieval systems [2].

REFERENCES

1. CODASYL Systems Committee, "Feature Analysis of Generalized Data Base Management Systems," CODASYL Systems Committee Technical Report (May 1971).

2. Martin, T. H., "A Feature Analysis of Interactive Retrieval Systems," Stanford University Report SU-COMM-ICR-74-1 (September 1974).

SERVING SPECIALIZED SUBJECT INTERESTS AMONG CHEMISTS

Lorrin R. Garson and Seldon W. Terrant, American Chemical Society, 1155 Sixteenth Street, N.W., Washington, D.C. 20036

ABSTRACT

Since early 1975, the Books and Journals Division of the American Chemical Society (ACS) has been investigating the potential of providing selected articles of interest to groups of chemists. These investigations have resulted in the *ACS CUSTOMIZED ARTICLE SERVICE* which was made available to ACS members starting in January 1977. This experimental service is offered in two subject areas: environmental chemistry and medicinal chemistry. Twice a month, subscribers are sent articles (on microfiche) selected from 18 ACS journals, along with the *ACS SINGLE ARTICLE ANNOUNCEMENT*; a subscription to the service also includes the *JOURNAL OF MEDICINAL CHEMISTRY* or *ENVIRONMENTAL SCIENCE & TECHNOLOGY*.

In January of this year (1977), the American Chemical Society (ACS) started offering the *ACS CUSTOMIZED ARTICLE SERVICE* on an experimental basis. The purpose of this new semi-monthly service is to assist chemists, and others, in covering the continually expanding journal literature by providing them with articles of interest in specific fields. The service was started after favorable reaction to several pilot experiments conducted earlier. Two subject areas are being covered: environmental chemistry and medicinal chemistry. Subscriptions are only available to ACS members at present.

The key feature of the service is a comprehensive selection of articles of interest to subscribers taken from all ACS journals except *CHEMICAL & ENGINEERING NEWS* and *CHEMISTRY*, a total of 18 journals. Subscribers receive the complete articles that are selected (on microfiche), plus the *ACS SINGLE ARTICLE ANNOUNCEMENT* (SAA) twice a month. SAA contains the tables of contents of the ACS journals that are currently published. In addition, subscribers are sent monthly issues of either *ENVIRONMENTAL SCIENCE & TECHNOLOGY* or the *JOURNAL OF MEDICINAL CHEMISTRY*, depending upon their choice of subject area.

In order to determine ACS members' interests, an extensive survey was conducted in 1974; questionnaires were sent to approximately 10,000 members, with 6500 responding. On the questionnaire, the eighty *CHEMICAL ABSTRACTS* sections were arranged alphabetically and chemists were asked to rank them according to their interests

from very interested, to not interested at all...in six levels. From these data the subject areas for experimentation were selected, namely: (a) environmental chemistry, (b) the chemistry of polymers and petroleum, and (c) medicinal chemistry.

There were a number of considerations in the selection of subject areas: (a) a reasonable number of people had to be interested in the subject, (b) a sufficient number of articles had to be published in the subject area to make the service attractive and useful, and (c) the articles had to appear in several journals. The latter requirement was necessary because if only one journal contained the great majority of the articles, an individual would be well served by subscribing to it.

It would be ideal to be able to provide individual chemists with all of the articles of interest to them from ACS journals, and only those articles. However, this would have to be accomplished in an economical and realistic manner. Scientists could be served in two ways: individually, or by clustering them into groups according to their common interests. In this experiment, we chose to serve groups of chemists rather than individuals.

PILOT EXPERIMENT IN ENVIRONMENTAL CHEMISTRY

In 1975, a group of 66 volunteers participated in a pilot experiment in environmental chemistry. They had indicated a high interest in the following *CHEMICAL ABSTRACTS* sections: (a) air pollution and industrial hygiene, (b) inorganic analytical chemistry, (c) sewage and wastes, (d) water, (e) organic analytical chemistry, and (f) toxicology. The participants were sent nine consecutive issues of SAA, copies of the articles selected, and a questionnaire with each mailing. In this initial experiment, our primary concern was to determine how well the articles were supplied as tear sheets (the form familiar to journal users) rather than on microfiche.

Two things were asked on the questionnaires: (a) of those articles sent, which were not of interest and should not have been included, and (b) of those articles cited in SAA, but were not sent, which should have been included. On the average, 78% of the participants returned the biweekly questionnaires. The average number of articles sent per issue was 20, and 75% of the respondents indicated interest in any article provided. The volunteers expressed interest in 19 articles per issue. The average recall and precision for article selection was 83% and 78%, respectively.

Many respondents indicated that they thought the service would be an excellent aid to practicing chemists and that the selection of articles had been very good. However, some criticized the service because of the exclusive use of ACS publications and felt that this was too limiting. At the conclusion of the experiment, each of the volunteers received a detailed questionnaire. A summary of the data is shown in Table I.

Table I: Participants' Responses to Pilot Project Environmental Sciences Group

Question	Form A* ($50/$250)	Form B* ($100/$250)
Would you subscribe?		
yes response	70%	64%
Which format?		
hard copy	5%	22%
microfiche	27%	33%
microfiche with ES&T	68%	44%
Access to microfiche reader		
readily accessible?	31%	
inconveniently accessible?	54%	
inaccessible?	15%	
Interested in purchasing a microfiche reader?		
yes response	65%	

* Two different forms of questionnaires were used. Half of the participants were asked if they would subscribe to the service with articles on microfiche at $50 per year, or articles provided on tearsheets at $250 (Form A). The other half were asked the same question with the prices set at $100 and $250, respectively (Form B).

The majority of volunteers indicated that they would subscribe to the service and would be interested in purchasing a microfiche reader. However, it would be naive to expect that such a proportion would actually subscribe or purchase a reader. It should be noted that the preference for hard copy over microfiche was strongly influenced when the price differential changed from 1:5 to 1:2.5. *ENVIRONMENTAL SCIENCE & TECHNOLOGY* (ES&T) consists of two parts. The first part contains reports, features, news items and various other sections; the second part consists of research papers. In this experiment, only the research papers were sent to the volunteers. As can be seen from the data, most of the volunteers would prefer the complete publication. A majority of the volunteers have access to a microfiche reader; however, somewhat more than half of them find access to readers inconvenient.

PILOT EXPERIMENT IN THE CHEMISTRY OF POLYMERS AND PETROLEUM

In this experiment 58 volunteers were sent nine consecutive issues of the service, as in the previous experiment. These participants had indicated a high interest in the following *CHEMICAL ABSTRACTS* sections: (a) elastomers, including natural rubber, (b) petroleum, petroleum derivatives and related products, (c) plastics manufacture and processing, (d) synthetic high polymers, (e) surface chemistry and colloids, and (f) organic analytical chemistry. However, one-third of the participants received articles on microfiche and the remainder received them as tearsheets.

As in the first experiment, each participant received a questionnaire with each issue and was asked to indicate whether the articles sent were of interest or which of the articles not sent should have been included. On the average, 73% of the participants returned the questionnaires. Each issue contained an average of 15 articles and 62% of the respondents indicated interest in any given article. The volunteers expressed interest in 15 articles per issue. The average recall and precision for article selection was 62% for both parameters.

The average number of articles received by participants in this group is somewhat lower than the number sent to participants in the environmental sciences group. Furthermore, the percent of respondents interested in any given article, the percent recall, and the percent precision, are also lower with the polymer/petroleum interest group.

These volunteers were also sent a questionnaire at the conclusion of the experiment. The data obtained are presented in Table II.

Comparison of the data in Tables I and II shows some significant differences. Of the participants in the polymer/petroleum interest group, considerably fewer indicated that they would subscribe to such a service, or would be interested in purchasing a microfiche reader. Also, fewer in the polymer/petroleum group have access to a microfiche reader. Although readers were less accessible to this group, none indicated they would want the service even in hard copy. The most frequently expressed complaint was that the articles published by the ACS are too theoretical to be useful to industrial chemists. Many in this group also mentioned that the exclusive use of ACS publications in the service is too limiting. Based upon experience with this pilot project, it was decided not to offer a *CUSTOMIZED ARTICLE SERVICE* in this subject area.

Table II: Participants' Responses to Pilot Project Polymer/Petroleum Interest Group

Question	Form A* ($50/$250)	Form B* ($100/$250)
Would you subscribe?		
yes response	25%	21%
Which format?		
hard copy	0%	0%
microfiche	75%	40%
microfiche and *MACROMOLECULES*	25%	60%
Access to microfiche reader		
readily accessible?	41%	
inconveniently accessible?	25%	
inaccessible?	34%	
Interested in purchasing a microfiche reader		
yes response	25%	

* Two different forms of questionnaires were used. Half of the participants were asked if they would subscribe to the service with articles on microfiche at $50 and $250 for articles provided on tearsheets (Form A). The other half were asked the same question with the prices set at $100 and $250, respectively (Form B).

PILOT EXPERIMENT IN MEDICINAL CHEMISTRY

In the spring of 1976, a pilot test was conducted involving 114 volunteers who had expressed interest in the following *CHEMICAL ABSTRACTS* sections: (a) pharmacodynamics, (b) toxicology, (c) general biochemistry, and (d) pharmaceuticals. From previous experience, it was concluded that providing articles on microfiche was preferable. Consequently, all the participants in this group were furnished articles on microfiche. To test the difference in acceptance of the service with only ACS journal articles, as opposed to articles from ACS journals plus others, permission was obtained from the American Pharmaceutical Association to use the *JOURNAL OF PHARMACEUTICAL SCIENCES* for this experiment. One-half of the participants received articles from ACS journals exclusively and the other half received articles from the *JOURNAL OF PHARMACEUTICAL SCIENCES* as well as ACS journals. As in the previous two experiments, each issue consisted of SAA, copies of articles from journals, and a questionnaire to determine which of the articles were of interest. However, with this pilot group articles were not taken from the *JOURNAL OF MEDICINAL CHEMISTRY*; the journal was provided in hard copy. Data obtained from questionnaires paralleled those obtained in the pilot study for environmental chemistry.

As in the other two experiments, a detailed questionnaire was sent to the participants at the end of the experiment. Approximately 80% of the respondents felt that including articles from the *JOURNAL OF PHARMACEUTICAL SCIENCES* would improve the service. Less than 10% of the respondents indicated that a microfiche reader was inaccessible, although about half said it was inconveniently accessible. Of those respondents who indicated that they would be interested in subscribing to the service, about 70% expressed interest in purchasing a microfiche reader.

Based upon these preliminary studies, the decision was made to introduce the *ACS CUSTOMIZED ARTICLE SERVICE* in January 1977. The service is available only to ACS members at $75 per subject area. For ACS members outside of the United States the cost is $100 per subject area. In our current experiment, subscribers are provided with SAA and copies of articles on microfiche solely from ACS journals. For the service in medicinal chemistry, it was decided not to include articles from the *JOURNAL OF PHARMACEUTICAL SCIENCES* during this first year. SAA and microfiche are sent approximately one week after the cover date of the respective issue of SAA; these two items are sent together by first class mail. The *JOURNAL OF MEDICINAL CHEMISTRY* and *ENVIRONMENTAL SCIENCE & TECHNOLOGY* are sent from the printer by second class mail. A hard copy subscription to the *JOURNAL OF MEDICINAL CHEMISTRY* or *ENVIRONMENTAL SCIENCE & TECHNOLOGY* is provided as part of the service. As an additional service to subscribers, arrangements have been made to provide microfiche readers at a discount. These readers are also available to subscribers of other ACS microfiche services.

Although it is too early to draw definitive conclusions from the on-going production experiment, preliminary data indicate that the service is being well received. The ratio of subscribers in environmental chemistry to medicinal chemistry is 2:1. The experiment will be continued in 1977 and, probably, throughout 1978. One or more additional subject areas may be added in 1978.

Excerpt From

CLASSIFICATION - HUMAN - VALUE OF

Peter Glickert
Chemical Classification Division, U.S. Patent and Trademark Office
Washington, D. C., 20231; (703) 557-1677*

KEYWORDS: Classification Operations; Classification Systems; Retrieval System Costs; User Effort; User Wants

ABSTRACT: I. Views of some authorities on the art of classification; II. Operations which humans can legitimately call "classification"; III. Requirements for a hypothetical retrieval system of maximum value to highly-paid users; IV. The value of the classifier; V. Costs involved in creating and maintaining an ideal system.

INTRODUCTION

This paper is purposefully designed to be inflammatory in the hope that a discussion of the material contained in it will produce both light and heat. The first section comprises an extensive series of quotations which have been juxtaposed to give the appearance of dialog. I hope the originators of these quotes will not be offended by any apparent mis-use of them.

In the second section, I extrapolate, mainly from the quotations in Section I, some other aspects of what can be legitimately be called "classification", so that further values inherent in such procedures may be recognized and/or debated. Using these expanded ideas, I sketch in Section III a retrieval system which is their culmination, and in Section IV I speculate on the value of the classifier, especially in the context of an ideal retrieval system, the costs and benefits of which are sketched in Section V.

II.B. EARLIEST EXAMPLE OF CLASSIFICATION

"...and God divided the light from the darkness. And God called the light Day, and the darkness he called Night. And the evening and the morning were the first day" (2).

Several items in this short story are noteworthy from a classification standpoint.

II.B.a. Classification can be practiced not only upon mental constructs but upon phenomena outside the human mind.

II.B.b. The story illustrates the two essential steps in the creation of a classification system: 1. Separating and gathering; 2. Labeling the categories. Classification *per se* involves the assignment of further items to the categories already established.

II.B.c. This is classy classification, in that it follows the spirit of classicism in its simplicity, restraint and proportion. It creates a simple two-value system, but like all such intellectual systems, it leaves cracks between the concepts. In

* Mailing address: 140 Twelfth St., S.E., Washington, D.C., 20003 The U.S. Patent and Trademark Office has nothing to do with the views expressed herein.

this case, are "evening" and "morning" to be categorized as Day or Night? Obviously an indexer and/or a searcher using such a system must have specific instructions for classifying and searching these "between-the-cracks" phenomena. And when the creator of a classification system prepares these instructions, will he or she be wise enough to extend the instructions to include sunrise, sunset, daybreak, dawn and sundown?

II.B.d. The word "day" is used in two different senses, but in the context of Genesis, "day" apparently is a category itself, which subsumes the classes Day and Night. If the classification schedule is arranged hierarchically, then "day" can be used as a generic term and as a residual category for species or varieties which would otherwise fall between the cracks of the specified species. One cannot do this on a computer, of course, which converts both day and Day to DAY.

II.B.e.v. If, in the schedule of Section II.B.d., additional symbols are applied to the classes, in the following fashion: 1 day; 1a Day; 1b Night; the following results will be observed:
1. The generic term is automatically "posted" when the specific term is "posted"
2. Where the search system permits variable-field search requests, the searcher has a choice between making a generic or a specific search, depending upon his need.

II.B.e.vi. If the system of II.B.d. also provides a facet for "almost", symbolized, perhaps, by an asterisk, then a searcher can ask for 1-*, to retrieve only the material concerned with the "between-the-cracks" phenomena of section II.B.c., whether they be classified originally as "almost Day" (1.a.*) or "almost Night" (1.b.*).

II.B.e.vii. When words are used to label categories, the labels are more meaningful to a searcher and classifier than are isolated arbitrary symbols; however, a faceted semantic code is more useful, that is, has a greater value.

II.C. LATEST EXAMPLE OF CLASSIFICATION

In writing this paper I have performed a lot of classification. I have formulated thoughts and applied words to these thoughts, that is, I assign these thoughts to categories which I believe to be represented by the words I use. Also, when the language provides "supraordinate" terms, I choose a level of specificity in the words I use (5). Thus, the formulation of verbal discourse involves the categorization of phenomena and the use of concept hierarchies internal to the mind of the formulator and perhaps, also, the categorization of phenomena external to the mind of the formulator.

II.D. WHEN DOCUMENTS ARE CLASSIFIED

II.D.a. Except for the identification of individuals (proper nouns), almost every expression used in a document constitutes a class preconceived by the author. Thus, an author assigns his document to a number of classes merely by writing it, but by including so many classes, this fact is obscured. When a number of documents are randomly collected together, the classes contained in the documents are further obscured.

II.D.b. A publisher does not publish every manuscript which is sent to him; rather, he selects publishable manuscripts on one or another preestablished basis and sends the others back to the authors. A journal editor, likewise, selects articles for publication on the basis of the "subjects" which his journal is designed to carry. A newspaper editor has different bases for choosing the stories he will carry.

II.D.c. All of these publishing and editing functions are classification functions and every data base from a single source is a classified data base.

II.D.d. Specialized libraries choose the material of interest to them, and this is also classification, making a specialized library a pre-classified data base. Every operating mechanized retrieval system discriminates as to what material will go into its data base, and also is thus pre-classified.

II.D.e. The kind of classification ordinarily thought of as classification happens when the documents are resorted on the basis of common threads contained therein. This occurs when a book is catalogued, when a paper is indexed or when a collection is searched, and all of these sorting operations tend to clarify the class membership of a document, originally assigned by the author when he wrote the material.

III. GOLDILOCKS AND THE THREE SYSTEMS

They say that President Carter, in selecting his cabinet asked the Chairman of IBM for information on who would be good choices. As a result, the present cabinet is said to bear no small resemblance to the IBM Board of Directors. Let us propose a different mode of information retrieval for him to use.

Public interest and consumer groups have long complained that policy-makers in Federal regulatory agencies frequently are hired from the industries which the agencies regulate or law firms which represent such industries, thus producing an industry bias in the agency. The agencies often defend themselves by explaining that it is those who worked in regulated industries who best know what it is that needs to be regulated. It may occur to some future president of the United States that a list of people previously associated with such industries or such law firms, but no longer working in them will be a good source for choosing unbiased but competent personnel.

(An indexed system is too hard to search for such personnel; a full "raw" text system is too soft.)

III.C. The System That Is Just Right

It is the size and organization of the data base which makes it just right: the main data base contains any and all published information in a form suitable for crude "data base refinement" before on-line iterative searching is to take place.

I have never doubted the possibility of a "World Brain". I have no doubt that a truly massive data base of abstracts, say, equivalent to the Library of Congress, can be stored and automatically searched without blowing all the electric fuses at the

information center. All of this data base does not need to be constantly spinning around. Once a firm, non-negotiable, demand is made on the computer industry to supply such hardware, industry will respond by providing such a system, but perhaps we the theoreticians should do the major design work.

The data base store will be substantially un-pre-classified, to retain accessibility and searchability; that is, the data base will consist of abstracts of the real information contained in the source documents, not just index terms taken from a list which is by its nature incomplete, and which rapidly grows obsolete. Vocabulary and language normalization, however, will be built into the system to substitute for classification to insure that the search question and the data base with which it is to be compared are isomorphic.

The application of the SMART system and its language processing principles to both the data base and the question is one way to assure one-to-one correspondence between the two. Even if the SMART system were to operate on a universal level, this system requires considerable processing for even a single sentence of the data base, and given the enormous size of the data base contemplated, the very slowness of the speed of light will prevent the processing of the data base from taking place overnight to produce a daughter data base. The data base itself must contain its information in a directly searchable format, free from the quirks of English grammar.

This format will be "topologically invariant", perhaps like Chinese grammar. While the crude "data base refinement" search will, as often as not, depend upon identification of a single "word" in the data base, the format is, of course, carried over to the daughter data base, where, in the more discriminatory later searches, the presence of language-like "roles" or "relational factors" can be important in improving the relevance of the ultimate search answer.

The "words" which comprise the data-base format are semantic, faceted codes, as described in Section II.B. It is likely that as many codes will need to be available as there are words in an unabridged dictionary. Each code represents a concept, rather than a mere word, and, indeed, often the facets which made up a code can be considered a rough "definition" of a concept. It is believed that any new words which come into natural languages merely express concepts which are a combination of old concepts and thus are accomodated by new combinations of old codes.

IV. THE VALUE OF THE CLASSIFIER

IV.A. Who Is A Classifier?

A person (or Deity) who performs any of the functions described in Sections II.B., II.C., II.D. or II.E.b., can be considered a classifier, that is, any human with the power of speech is a classifier. Of those who do classifying for a living, however, we may consider the person who sets up a classified schedule as a professional classifier, while those who perform the technician functions of assigning further items to the schedule are indexers or cataloguers.

IV.B.c. The Threat Of The Machine

There can be two kinds of "automatic indexing". In one, the machine will perform the almost clerical task of assigning documents to previously arrived-at index terms. But advocates of full automation may well insist that the setting up of the categories themselves can be handled by machines on a statistical basis, thus devaluing the role of the professional classifier, as well as the indexer.

A decision to automate indexing tends to leave out the classifier, who is concerned with what the word and the work mean; who detests statistics because of the intellectual vacuum in which they operate. Which will win, the mathematics or the semantics? Many classifiers may already have decided the winner and joined the mathematicians in praise of automatic indexing. After all, the results of a fully automatic system may be only slightly more inadequate than a hand-indexed system. I believe that the better road for classifiers is toward more semantics and away from indexing itself.

IV.B.d. The Role Of The Classifier In The Ideal System

It is professional classifiers, of course, who are best qualified to work out the notations for the code dictionary. Their experience with word meanings and the semantic relations between words fits them ideally to do this work. Basic conceptual terminology must first be arrived at, but the code dictionary need not be done in alphabetical order.

Let us consider first the enormous and time-consuming task which a searcher has in the mere operation of finding suitable data bases to search and in selecting the one he hopes he will get his money's worth from. If the first task of the code dictionary makers is that of codifying the words and expressions used as index terms in all currently available systems, this will take care of a large, and perhaps the most important, part of the dictionary. Thus our World Brain can spend its infancy as a master file of all currently available indexed systems and spare the searcher the need for manual data base refinement.

If, for one reason or another, new indexed systems are created, classifiers need only select from the dictionary those specific terms which are appropriate to the system, the hierarchy and thesaurus arrangements being taken care of inherently by the facets of the specific codes.

It remains to be seen, further, whether devising the "topologically invariant" format for the information language is properly the domain of the linguist or the classifier. The former have devoted themselves almost exclusively to the "well-formedness" of expressions rather than their meaning content. Thus, assigning semantic tasks to linguists is something like assigning a French Maitre d' to the job of dietician, that is, shifting the emphasis of his job from table manners to the nutritive value of the food being served.

V. MONEY

The ideal blue-sky system outlined in Section III.C., is

designed, of course, for users of the system to live happily ever after: to find what information has been published -- no matter where, no matter when -- relative to the user's needs. In particular, the system is designed to spare the user the cost of low relevance, this cost being the time he must spend looking through a pile of documents of marginal utility, the pile increasing perhaps exponentially in size with each percentage point of increased recall. Everyone can recognize the social disutility of wasting the time of a President of the United States who likes to hold her cards close to her chest. However, Ph.D. research scientists and patent examiners also are paid quite well for their time and both try to avoid, as much as possible, having their literature searches filtered through information specialists.

The user is living happily ever after because the system shifts the cost to the developer, creator and operator. Obviously, the processing of entire natural language abstracts into a machine language format and vocabulary will be more expensive than merely indexing them, but is it an impossible expense? Hopefully, enough of the work of data base creation can eventually be routinized that machines will do most of the work. At first, however, the process will be labor intensive. This in itself is no reason to curl up our toes and say, beforehand, that these blue skies are too expensive. America does not do without dolls and color television sets because their manufacture is too labor-intensive; rather, manufacturers look to the Far East, where labor is abundant and cheap to have these things done. I want to point out that India has a perennial unemployment problem among college graduates who are fluent in English.

V.B. Operating Expenses

It is expected that a single information center or a unified system of such centers will operate the massive parent data base, and that this data base will be searched serially, but in batchwise fashion to reduce costs. Where the "subject" of the data base refinement search is of widespread interest, such as "Microorganisms affecting human liver functions" the daughter tape may be retained on a semi-permanent basis for repeated on-line searching, by many users, being erased and recreated only when up-dating is called for. The daughter data base will probably be much smaller than data bases now available for on-line iterative searching, and will reduce on-line computer costs.

CONCLUSION

Classification procedures are present in almost all human activities, especially those which involve words. However, in information retrieval involving mechanized systems, classification of entire documents before storage has a negative value; rather, classification, based upon a search question, and using symbol equivalence, is the most universal answer to search problems. The development of a language for storing information which is susceptible to such searching, is a job which will ensure the productive employment of professional classifiers for many years to come.

THE NASA TECH BRIEF PROGRAM:
A COST BENEFIT EVALUATION

by: F. Douglas Johnson
Transfer Research and Impact Studies Project
Industrial Economics Division
University of Denver Research Institute
Denver, Colorado 80208

(Cost Benefit Evaluation, Economic Measures, Statistical Methods, Technology Transfer, Value of Information)

NASA initiated the Tech Brief Program in 1963 to disseminate selected technical innovations. DRI, under NASA contract, accumulated data since 1968 for more than 120,000 document requests from individuals in all sizes and types of organizations. This data bank was used to conduct a detailed cost benefit evaluation of program activity between 1971 and mid-1976. The methodology included: detailed NASA cost analysis; a two-tiered random sample of requesters; telephone interviews; statistical analysis of requester costs and benefits; and derivation of benefit modes and expected net values. Mode 1 benefits relate to document availability and content review. Mode 2 is applying document content to modify existing practice and Mode 3 relates to new practices. A net benefit probability function was derived for each mode and these were combined to give the benefit function for an STI transaction through this program. The net user benefit to NASA cost ratio was 11:1. Evaluation costs were less than five percent of program costs.

The data are being analyzed to discover relationships among benefit, requester, and technology variables so that operational strategies can be designed to decrease NASA costs and increase requester net benefits. Strategy options include requester selection, NASA selection and Tech Brief marketing. If this methodology can be applied by others to characterize benefits for STI systems, it may become possible to estimate the potential value of a technical document by predicting the potential user population and expected benefits.

Mishan, Edward J. Cost-benefit analysis. NY: Praeger Publishers, 1976.

Robertson, S.E., R. Reynolds, and A.P. Wilkin, "Standard costing for information systems: background to a current study." Aslib Proceedings, 22:9 (September 1970) 452-457.

National Science Foundation. Office of National R&D Assessment. Federal technology transfer: an analysis of current program characteristics and practices; executive summary. Washington, DC: December 1975. (NSF 76-401; NTIS: PB 253-105)

BASIC ISSUES IN DESIGNING A PRICING STRUCTURE FOR LIBRARY AND INFORMATION SERVICES

Thomas R. Kochtanek, Case Western Reserve Univ., Cleveland, Ohio

In the near future certain forms of information, especially those provided as a service, are likely to be coupled directly with user fees for service. The question thus arises as to the basis for determination of user fees for services rendered. The purpose of this paper is to explore that basis and to suggest possible analytical tools which might aid in the design and implementation of a "pricing structure" for such services.

Why a charge for services rendered? Basically, the argument is as follows. In most libraries and information centers (hereafter referred to as systems) there are costs which must be financed. Revenues to cover these costs can be generated by allocating fees to certain services. Normally revenues are generated indirectly, but in the cases we will consider (certain services; certain systems) revenues will be directly generated via a pricing structure. The logic is that those who benefit should pay. In this context direct users are considered as primary beneficiaries, and the charges they accrue are considered as the revenues which support the facilities offering those services.

Just what is a pricing structure? Why are they necessary? A pricing structure is simply a methodology for matching specific services with specific costs. Services and costs will vary from system to system, but the methodology remains essentially the same. The necessity of such a structure is seen when exploring the transmission of information. Consider the scientific/technological areas of research and development. Whether the user be an individual scientist/engineer, an industrial or academic research establishment, a technical society, or a government agency the need for the effective transmission of information is evident. The purpose of a system such as the library or information center is to facilitate this effective transfer. A pricing structure is one link between the system and the user, with the explicit purpose of generating revenues which allow the system to continue operating effectively.

So far we have used the terms "system", "service", and so forth in a rather vague manner. Let us attempt to clear these matters up by discussing the basic issues involved in designing a pricing structure for a library or information center. The basic issues are:

I. Types of Systems
II. Types of Services
III. Measurement of Service
IV. Value of Information
V. Fairness

I. Types of Systems Systems offering information services can generally be categorized according to their respective objectives and affiliations. The objectives of a system might be to:

A. maximize service
B. maximize profits
C. variations of A and B (in combination)

The design of a pricing structure necessarily involves the formulation of an objective function, the quantifiable version of the objectives. It must be stressed that the objective functions of these systems are indeed complex constructs. This is due to the fact that the objectives of systems such as these are not directly economic; many intangibles are inherent.

A second categorization is affiliation, defined as the coupling of two or more systems for the expressed purpose of providing increased or improved services. Typical affiliations are with governmental, industrial, academic, or private research corporations.

II. Types of Services Many types of information services will exist depending upon the nature of the supplier and the demand of the user. Some general categories are noted here:

A. current awareness systems
B. state of the art reports
C. problem solving
D. fact finding
E. bibliographic compilation
F. document retreival
G. inter-system loans of documents

Certainly this is not an exhaustive list. Specific services exist and will continue to be developed, and these services vary with the systems which offer them. An important consideration is the fact that levels of service can be offered to users who wish to specify the depth of a desired service. Another consideration is timing. The delivery of a particular service is an important aspect of that service, and the existence of various current awareness systems bears this out. Thus the time of delivery of a particular service is interwoven with the cost and the value of that service.

III. Measurement of Service A quantification of various services and their respective levels needs to be developed. The question of "how good is the service?" needs to be expanded from its present qualitative state to a state which measures the effectiveness of that service.[1] This measure needs to be defined in terms of the user, which is precisely why quantitative measures of the effectiveness of services are difficult to formulate. They are dependent upon the nature of the user who determines

A. what service is (how that user is served)
B. the usefulness of that service (validity)
C. the gain to different users

A body of literature is being developed which focuses on various methodologies for the measurement of library and information center services.[2,3,4,5]

Literature analyzing services speak mainly in a qualitative voice. The analyses are superficial; they often discuss various options, rarely quantify them.[6] There is a need for definitions of service in operational terms.

IV. Value of Information Information is a slippery concept. It defies all but technical representation. Is information a commodity, as is the concept of "goods" in economics? Does information submit to economic analysis? If so, what are its properties? This would be dependent upon several notions:

A. prior knowledge of the user
B. relevance of the supplied information to that particular user
C. the situation (time, place, context)
D. availability; accessibility

It has been argued that information does not submit to economic or value analysis, since it is not a material or consummable item.[7] Perhaps information plays a role similar to that of the catalyst in the chemical reaction. A catalyst stimulates or triggers a change, yet is not directly consumed in the ensuing reaction. The question currently remains unanswered: how can value be assessed to information?

V. Fairness The question of fairness pertains to the relationship between the supplier of information services and the user of those services. Thus fairness has two sides. From the supplier's point of view fairness of input should be enacted. This means that the user is charged (according to the pricing structure) for every request or query for which a service is provided. The user may reasonably expect fairness of output, meaning he pays only for those services which he judges to be "useful". This implies a subjective judgement by each user, which may or may not be immediately assessable. Considerable effort needs to be put into this area.

Having discussed the basic issues in designing a pricing structure we will now turn our attention to the hypothetical design of the structure. In the discussion we will consider the Operations Research approach as a possible methodology for design.

The initial step in the methodology involves defining the objectives of the library or information center. Objectives, objective functions and their general characteristics were briefly touched upon in section I Types of Systems. The next step would be to flowchart, in detail, all processes which form that system offering the service(s). Included would be the various requests for service and the specific mechanics of the search and/or preparation needed to fill those requests. The result would be a functional and physical description of how the system operates to provide its services.

The identification of alternative courses of providing the desired service would follow. This is an important note and should be considered seriously. The enumeration provides the analyst with several courses of action, an important feature when one considers the multiplicity of forms a particular request for service may take. If all requests were roughly equivalent a single algorithm could be set up to handle all requests. Considering the number of requests, users, services, and systems offering services, the analyst should be capable of designing alternative courses of action for providing the desired service. But a measure of performance must be identified which would be capable of evaluating alternative courses of action with respect to the probabilities of achieving the desired result.

Following the formulation of the performance measure, an attempt would be made to describe each process or operation with a mathematical expression. The variables in the expressions would necessarily reflect those measures of cost and of service, since these are the variables of the objective function. This is the are where the measurements of costs and service are so crucial. Unfortunately, they are difficult to formulate (see section III). This is the weak spot of this particular methodology. Services and costs must be quantified for the processes or operations of the system providing that service. These expressions must then be integrated into a single mathematical model of the service operations. The analyst then uses the model to optimize the services and costs so as to attain the highest proportion of each stated objective. This solution could then be implemented and evaluated to test its performance with respect to the objectives. Changes would be made if necessary, and their consequences would be evaluated by the same performance measure. If the model proves to be promising it can be generalized to accomodate other systems and services.

Thomas R Kochtanek
Case Western Reserve University
Cleveland, Ohio

1. Orr, Richard H. "Measuring the Goodness of Library Services: a General Framework for Considering Quantitative Measures." Journal of Documentation, 29:1 (September 1973) p.315-332.

2. Kantor, P.B. (personal communication).

3. Kantor, P.B. "The Library as an Information Utility in the University Context: Evaluation and Measurement of Service." JASIS, 27:2 (March-April 1976) p.100-112.

4. Ackoff, R.L. and Emery, J.C. Third Version of an Idealized Design of a Scientific Communication and Technology Tranfer Systems (SCATT). Pittsburgh, Pa. Busch Center. 1974.

5. Hamburg, M.; Clelland, M.C.; Bommer, M.W.; Ramist, L.E.; Whitfield, P.M. Library Planning and Decision Making Systems, MIT Press: 1974.

6. Wish ,J.R. and Wish, M. "Marketing and Pricing of Online Services" ASIS Midyear meeting proceedings. May 15-17, 1975. Wash.,D.C. p.143-158.

7. Swanson, Rowena Weiss. "Design and Evaluation of Information Systems." in ARIST v. 10 (1975) ed. by Carlos A Cuadra. p. 43-101.

COSTS OF ON-LINE BIBLIOGRAPHIC SEARCHES IN INDUSTRY

P.A. Komoroske and M.F. King
General Electric Company
Whitney Library
P.O. Box 8
Schenectady, NY 12301

A survey of on-line costs, time and data bases used during one years searching at the General Electric Research and Development Center in Schenectady, New York has been compiled. Charts showing costs of individual searches for authors, for subjects, and average costs for all searches per month are included. Data is based on the use of the on-line systems available through Systems Development Corporation and Lockheed and cover the period from October 1975 through December 1976.

Bibliographic searches - costs

COSTS OF ON-LINE BIBLIOGRAPHIC SEARCHES IN INDUSTRY

P.A. Komoroske and M.F. King
General Electric Company
Whitney Library
P.O. Box 8
Schenectady, NY 12301

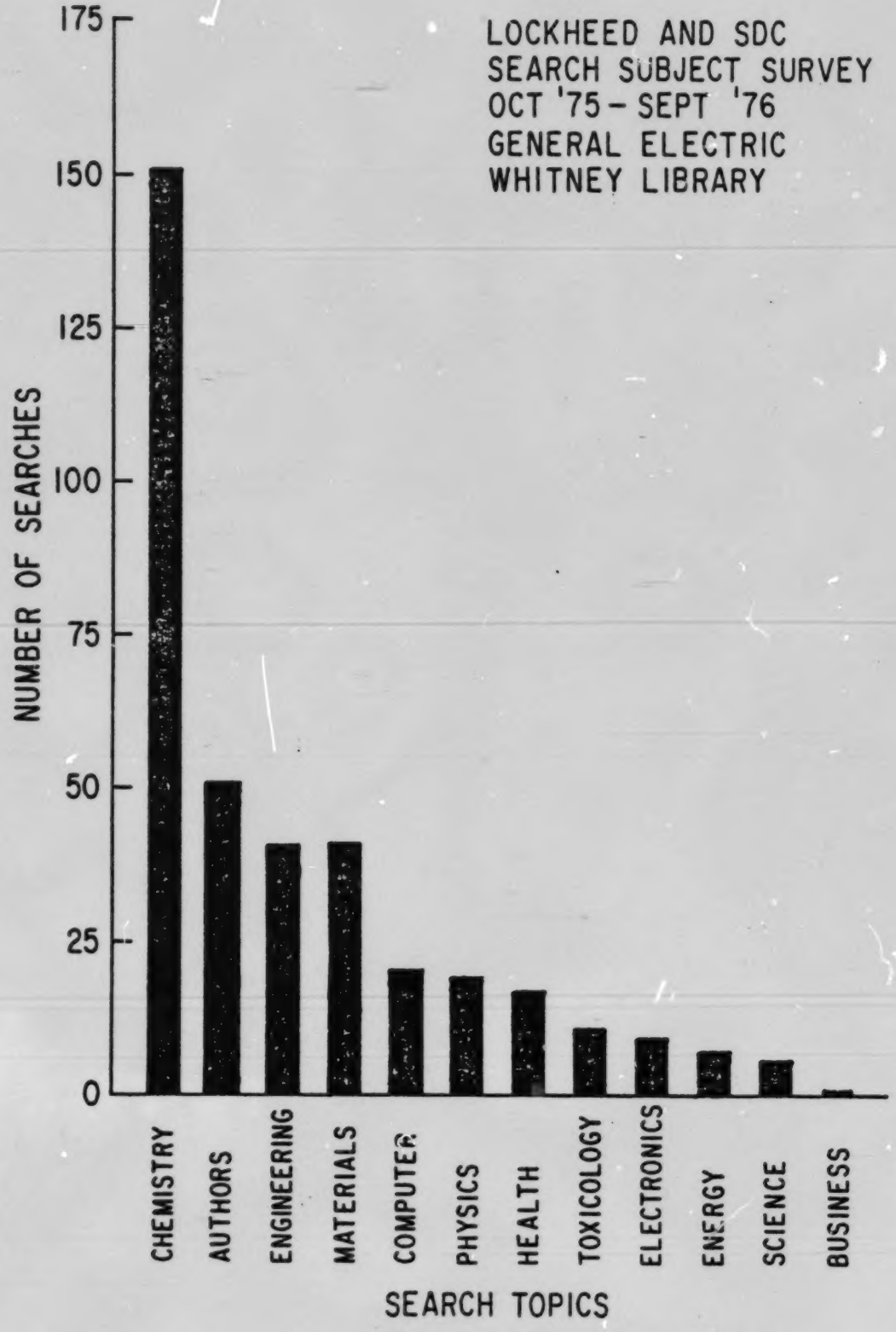
Through the General Electric Whitney Library, the over five hundred staff members at the Research and Development Center and General Electric personnel nationwide may obtain on-line search service. Both the Lockheed DIALOG and the System Development Corporation retrieval systems are used. Average costs for all searches per month and costs for individual searches were analyzed for the period October, 1975 through September, 1976. Additional data concerning DERWENT patent searches on SDC will also be duscussed. Library reference work, user demonstrations, and on-line current awareness search statistics are not included. Costs of individual searches are comprised of search and telecommunications connect fees besides a fifteen percent library service surcharge.

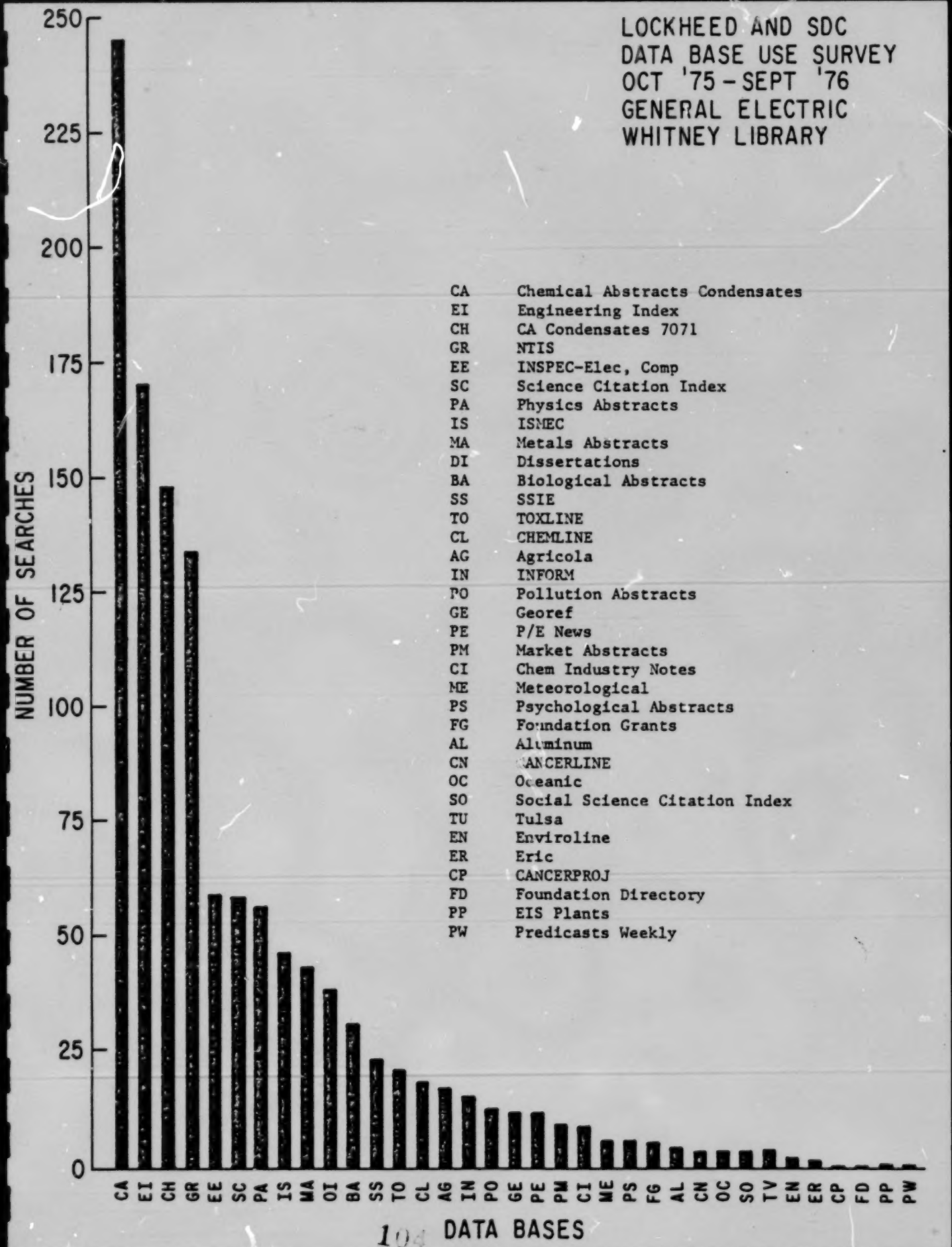
As background for the cost information, a subject search survey is included (Figure 1). The data on the chart also incorporates searches done on the National Library of Medicine TOXLINE system. Author searches include development of bibliographies and investigations of cited references or corporate source listings. The high proportion of chemistry orientated searches reflects the interdisciplinary and international range of CA Condensates data base and of the interests of the research staff. Most scientists prefer to use the on-line systems to search the weekly issues retrospectively. The over eighty dollars per connect hour energy and business data bases are not used extensively. The data base use survey chart (Figure 2) complements the subject survey and also illustrates that the low cost ERIC data base is searched only occasionally.

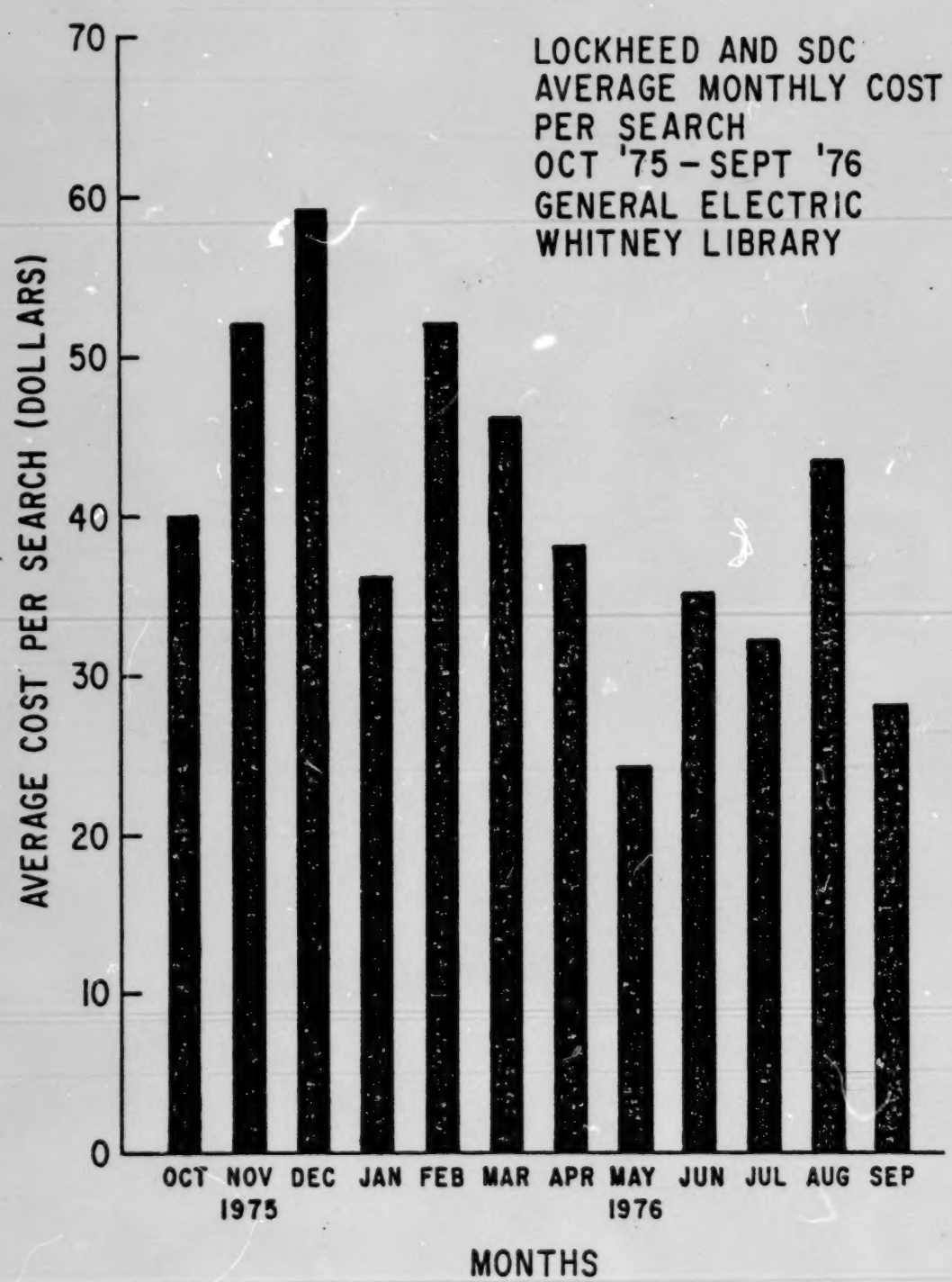
The average of the monthly cost figures given in Figure 3 is $840.41. This amount corresponds well with the thirty to fifty dollar cost estimate usually quoted during the user interview at the Whitney Library. Searches are highly responsive to need in this industrial institution so no limit is placed if all on-line printing is requested.

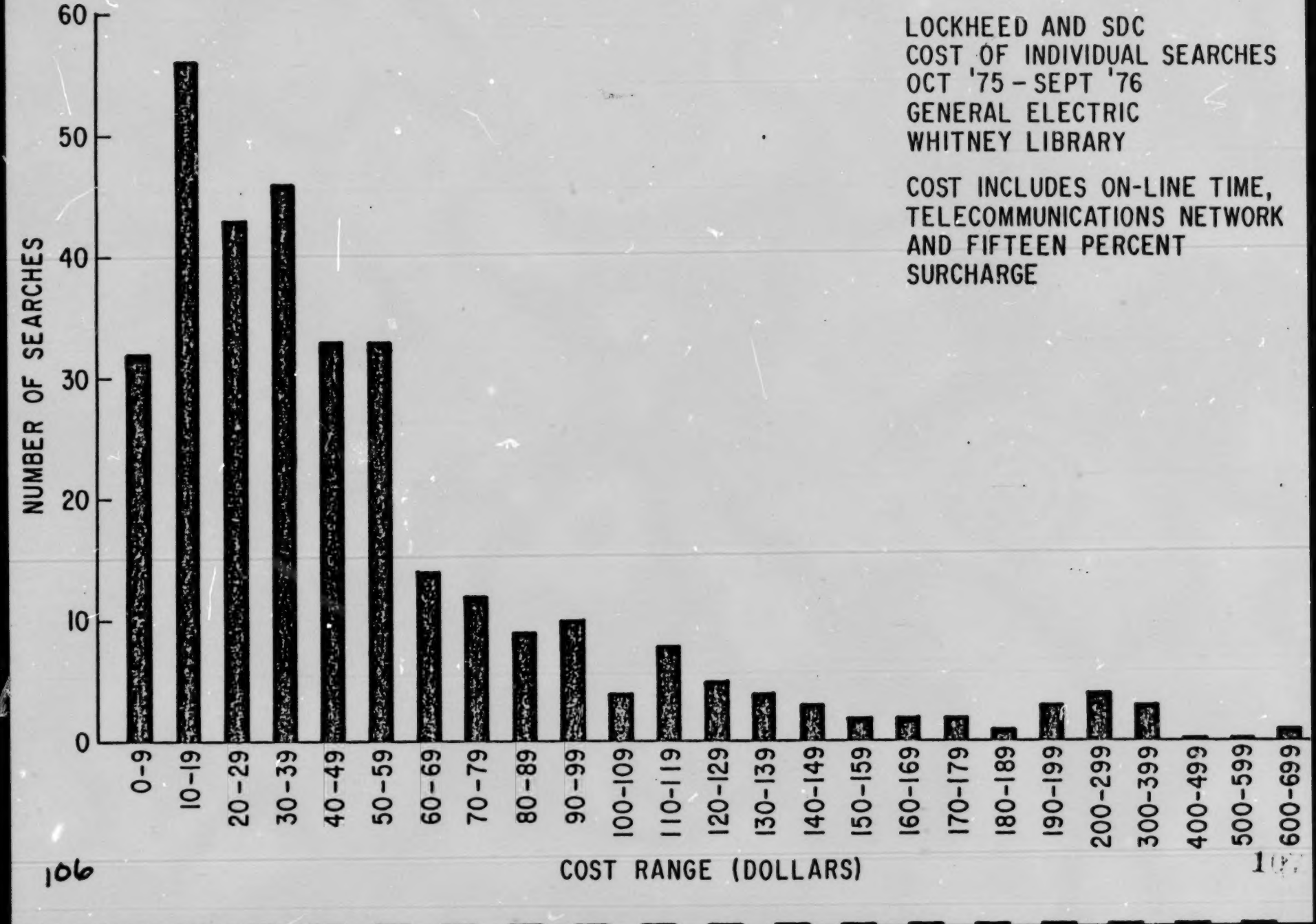
Costs for individual searches actually varied from $3.22 to $677.38. A predominant number occurred in the ten to nineteen dollar range. Individual cost figures were rounded to the nearest dollar (Figure 4).

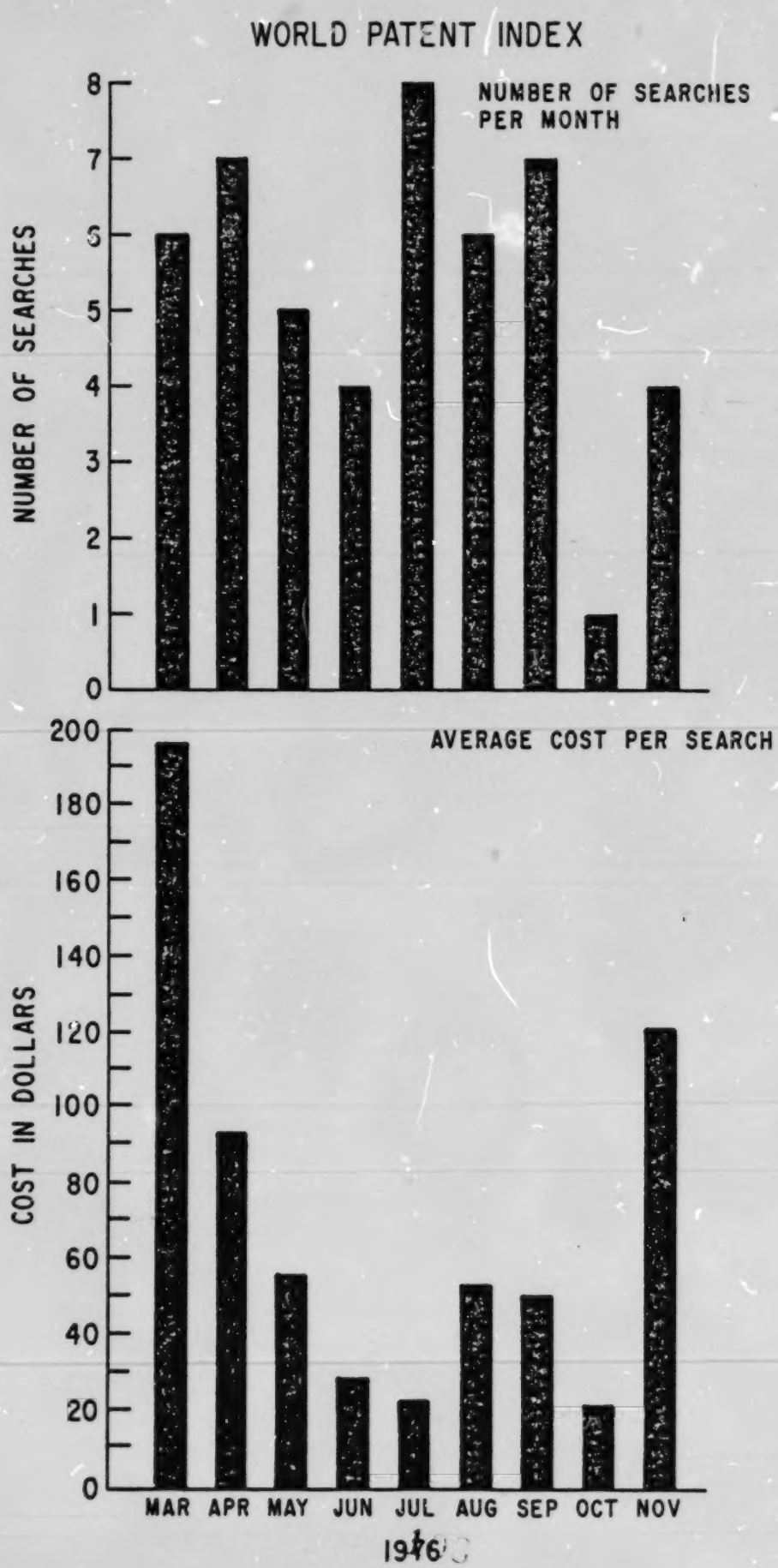
The Whitney Library also uses an on-line file for searching the patent literature. This file is available by special subscription arrangement through the System Development Corporation. It is a worldwide patent file that may be accessed by patent number, subject and company code. Since a patent file is not used in the same way that a bibliographic file is used, the number of searches and their costs show a more variant pattern. This file has been available since March 1976 so a full 12 months statistics are not available yet. However, a trend can be shown for both usage and average costs. The average number of searches performed each month is 5 with an average cost per search of $70.33. This special file fills a specialized need for research and development organizations. It is a unique file with many special indexing features supplied by the Derwent Publications staff. Costs for searches shown here include only the on-line charges telecommunications, off-line citation print costs and a fifteen percent library surcharge (Figure 5).


LOCKHEED AND SDC
SEARCH SUBJECT SURVEY
OCT '75 - SEPT '76
GENERAL ELECTRIC
WHITNEY LIBRARY
NUMBER OF SEARCHES
175
150
125
100
75
50
25
0
CHEMISTRY
AUTHORS
ENGINEERING
MATERIALS
COMPUTER
PHYSICS
HEALTH
TOXICOLOGY
ELECTRONICS
ENERGY
SCIENCE
BUSINESS
SEARCH TOPICS


LOCKHEED AND SDC
DATA BASE USE SURVEY
OCT '75 - SEPT '76
GENERAL ELECTRIC
WHITNEY LIBRARY
NUMBER OF SEARCHES
250
225
200
175
150
125
100
75
50
25
0
CA Chemical Abstracts Condensates
EI Engineering Index
CH CA Condensates 7071
GR NTIS
EE INSPEC-Elec, Comp
SC Science Citation Index
PA Physics Abstracts
IS ISMEC
MA Metals Abstracts
DI Dissertations
BA Biological Abstracts
SS SSIE
TO TOXLINE
CL CHEMLINE
AG Agricola
IN INFORM
PO Pollution Abstracts
GE Georef
PE P/E News
PM Market Abstracts
CI Chem Industry Notes
ME Meteorological
PS Psychological Abstracts
FG Foundation Grants
AL Aluminum
CN CANCERLINE
OC Oceanic
SO Social Science Citation Index
TU Tulsa
EN Enviroline
ER Eric
CP CANCERPROJ
FD Foundation Directory
PP EIS Plants
PW Predicasts Weekly
CA EI CH GR EE SC PA IS MA OI BA SS TO CL AG IN PO GE PE PM CI ME PS FG AL CN OC SO TV EN ER CP FD PP PW
DATA BASES


LOCKHEED AND SDC
AVERAGE MONTHLY COST
PER SEARCH
OCT '75 – SEPT '76
GENERAL ELECTRIC
WHITNEY LIBRARY
AVERAGE COST PER SEARCH (DOLLARS)
70
60
50
40
30
20
10
0
OCT
NOV
DEC
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
1975
1976
MONTHS


LOCKHEED AND SDC
COST OF INDIVIDUAL SEARCHES
OCT '75 - SEPT '76
GENERAL ELECTRIC
WHITNEY LIBRARY
COST INCLUDES ON-LINE TIME, TELECOMMUNICATIONS NETWORK AND FIFTEEN PERCENT SURCHARGE
NUMBER OF SEARCHES
0
10
20
30
40
50
60
0-9
10-19
20-29
30-39
40-49
50-59
60-69
70-79
80-89
90-99
100-109
110-119
120-129
130-139
140-149
150-159
160-169
170-179
180-189
190-199
200-299
300-399
400-499
500-599
600-699
COST RANGE (DOLLARS)


WORLD PATENT INDEX
NUMBER OF SEARCHES PER MONTH
NUMBER OF SEARCHES
0
1
2
3
4
5
6
7
8
AVERAGE COST PER SEARCH
COST IN DOLLARS
0
20
40
60
80
100
120
140
160
180
200
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
1976

"Accessing Numeric Data via Flags and Tags: Progress Report on a Real-World Experiment"

by

James P. Kottenstette and James E. Freeman*
Denver Research Institute
University of Denver

The Scientific and Technical Information Office, with financial support from the National Science Foundation, is conducting a data tagging and flagging experiment to evaluate the usefulness of Data Summaries and Data Terms in the NASA scientific and technical information system. The University of Denver, Denver Research Institute (DRI), is providing design and evaluation assistance on this experiment which will be completed in September 1977.

DATA SUMMARIES AND DATA TERMS

The principal objectives of this prototype project are: first to extend the concepts of data flagging and tagging to the aerospace scientific and technical literature; second, to obtain experience with the assignment of data flags and tags by documentation specialists; and third, to obtain "real-world" assessments of the costs and benefits of including data flags and tags in information products and services. The flagging and tagging employed follows the concepts suggested at the Energy R&D Data Workshop held at the National Bureau of Standards in 1974—concepts which are adapted to provide a DATA SUMMARY for accessions appearing in the IAA and STAR Abstract Journals.

The DATA SUMMARY is conceived as a free-text field which forms an integral part of the accessions in six subject categories appearing in the citation portion of the Abstract Journals. The DATA SUMMARY (DS) field follows the abstract in the familiar accession format and it is composed of three elements, presented sequentially, after the phrase DATA SUMMARY appears in the accession. The DS elements are: (1) a narrative phrase, clause or sentence that qualifies the data appearing in a journal article or a report, e.g., "a performance comparison of five resin systems"; (2) a string of terms representing the major variables for which numeric data are presented in the document; and (3) an explicit review of the numeric data representations (i.e., figures and tables) appearing in the document. Thus, the narrative phrase (element 1) equates with the Data Flag, and the list of terms representing major variables for which numeric data are reported (element 2) equates with the Data Tag.

---

*Messrs. Kottenstette and Freeman are co-managers of the Institute's Knowledge Utilization Program.

The assumption is that the content of the DATA SUMMARY should help the user make a more informed decision about the potential value of the document cited or accessed. Thus, for purposes of the experiment, the guiding principle in the construction of the DATA SUMMARY is to specify the manifest data contents of a document. Since the format, treatment, and contents of articles and reports are highly variable, the procedures and guidelines for the construction of data summaries have the objective of facilitating the formulation of adequate and consistent data summaries. That is, two different abstractors would reach essentially the same conclusions about the content of a DS.

The Data Tagging and Flagging Experiment results in data summaries appearing in STAR and IAA issues 05-10 (1977). Approximately 1500 accessions will appear in six selected categories (Aerodynamics 02; Composite Materials 24; Metallic Materials 26; Nonmetallic Materials 27; Acoustics 71; and Optics 74). Four types of data summaries are prepared, based on the data characteristics of the original documents: 1) Documents containing no numeric data; 2) documents containing excessive numeric data tabulations; 3) documents containing the usual data content; 4) documents for which it is impossible to develop valid data summaries.

A photocopy of a Type 3 DATA SUMMARY (the most common type prepared), taken from IAA issue 05, appears below.

c05-26

A77-16773 // Carbon diffusivity in iron-chromium alloys. K. E. Blazek (United States Steel Corp., Monroeville, Pa.) and J. R. Cost (Purdue University, Lafayette, Ind.). *Japan Institute of Metals, Transactions*, vol. 17, Oct. 1976, p. 630-636. 24 refs. NSF Grant No. GH-33574; Grant No. DAHC-0213.

Isothermal diffusion of carbon in ferritic and austenitic iron-chromium alloys has been investigated as a function of chromium concentrations ranging from 0 to 3.5 at% using C-14 and conventional sectioning and counting techniques. Measurements at four temperatures from 716 C to 1082 C revealed that the diffusion coefficient for carbon decreases roughly 10% per atomic percent of chromium concentration when a linear variation is assumed. A nonlinear analysis shows a small maximum near 1 at% chromium in the austenitic alloys. These results are discussed in terms of a simple model for trapping of carbon atoms by chromium. (Author)
DATA SUMMARY: Eight alloys, carburized in C-14 methane and homogenation annealed, were tested at 716 C, 963 C, 968 C and 1082 C; variables are chromium concentration, carbon concentration before carburization, impurity content, carbon-14 profile, inverse error function of carbon activity, carbon diffusivity, temperature, trap size, critical trap concentration, and interstitial sites in trap; four tables and five figures include numeric data.

This format is also used for entries appearing in the NASA/RECON system, to facilitate computer based retrieval. In addition, Data Terms appear in the subject indexes to STAR and IAA for these data-tagged documents. A Data Term is the NASA Thesaurus term that has been assigned to a variable or parameter on which quantitative measures are reported in the document. The Data Terms have no identification symbol in the subject indexes but are labeled DT in the NASA/RECON entry display. These data terms will be retrievable on NASA/RECON as regular subject terms or separately as data terms by using DT/(term).

Hypotheses to be Tested

The evaluation plan includes a test of three hypotheses: there should be a user preference for documents that have data summaries; there should be a decrease in the search time for users to actually reach relevant documents; there should be an increase in the precision factor for documents provided in literature searches and printouts.

The relevant variables have been identified as (1) users, (2) user preference, (3) search time, and (4) precision factor. These variables are defined as follows:

Variable 1: Users. Users are persons currently using NASA's access journals--STAR and IAA--and/or persons currently using NASA/RECON.

Variable 2: User Preference. User preference is defined as the extent to which users choose to use data terms and data summaries.

Variable 3: Search Time. Search time is the elapsed time for the combination of events and processes associated with a user's attempt--either alone or with others' assistance--to obtain information or data presented in documents. For purposes of this experiment, only those searches involving the use of STAR, IAA, and/or NASA/RECON will be considered.

Variable 4: Precision Factor. The precision factor is defined as the specificity of a user's decision concerning the probable numeric data contents of referenced items.

The data required to test the hypotheses is being derived from both user and intermediary document retrieval and ordering practice. The American Institute of Aeronautics and Astronautics, the NASA Scientific and Technical Information Facility, and organizations with standing orders for STAR and IAA documents (including NASA Centers) are participating in the data acquisition process, particularly in the identification of individual users for interview purposes.

The Marketing Plan

In many respects the experiment would be superficial if the need for systematic marketing of the concept of data summaries was ignored. The approach purposely has a "low key" character. It is designed to reach the potential user community in such a way that the user's document ordering behavior will not be affected unless s/he finds that data terms and data summaries are useful.

Objectives. The marketing objectives are limited strictly to:

1. Creating awareness in the user community that NASA is testing a particular technique for indicating the numeric data contents of aerospace literature; and

2. Familiarizing the user community with factors related to the use of data terms and data summaries.

Mechanisms. Two formal mechanisms are now used to achieve these marketing objectives:

- Publishing a "Special Notice" in each edition of STAR and IAA that contains data terms and data summaries; and

- Distributing a "Notice of an Administrative Experiment" to persons requesting documents from AIAA and STIF during the experimental time period.

By achieving these objectives, it will be possible to make a fair test of the usefulness of data terms and data summaries in the real world aerospace environment.

?BENEFITS OF INFORMATION

Marilyn M. Levine
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

Benefit-cost analysis as an economic tool arose in the 1930's as a public policy evaluation technique linked historically with the name of Vilfredo Pareto. The concept of "what counts as a benefit", known as the Pareto Criterion, holds that a change is desirable (a benefit) if it would make some persons better off (in their own judgment). Conversely, a change is undesirable if it would make some persons feel worse off and no one feel better off. Unfortunately, such neat categories are rarely found in the real world. To quote an old maxim, "It's an ill wind that doesn't blow somebody some good." What we must look at, therefore, is the relative mix of desirable and undesirable consequences, hoping to maximize the desirable ones.

Because of the subjective nature of the Pareto Criterion, economic benefit-cost analysis, closely allied with cost-effectiveness, is now considered "qualitative evaluation of alternatives"(1). The problem with applying this approach to information is that the first step in any evaluation technique involves choosing appropriate indicators of success, or something to maximize.

What would the benefits of information look like? One suggestion has been made (2) that people who play the information game well have power. If they have power, how much information does it take to gain such power? How much social power do they have? Can one measure social power by the amount of money or the number of people one can command given a certain amount of information? These questions will be addressed.

Another suggestion often made in connection with information is that a gain in information leads to a bettering of decision-making. The more information, the better the decision. The benefit, in this case, is the avoidance of costly mistakes. If the federal government, for example, could have known in advance that there would be problems with the Pruitt-Igo housing project, it would never have been built.

Prediction of events can also be a benefit of having information. This is exemplified by two sorts of situations. In one, such as disease or earthquake prediction, the benefit of having the information can be measured in terms of lives saved. In the other, such as stock market or horse-race predictions, the benefit of knowing what is about to happen is worth whatever the market will bear. (Theoretically, this kind of benefit is unbounded.)

Prediction goes hand-in-hand with control, so that if one wants to control a situation to one's own benefit, one must be willing to pay the cost of the information. This also gets us back to the social power situation, where prediction and control are power concomitants.

It thus appears that the benefits of having information can be likened to having the right tool to do the job. Metaphorically, if you want to drive a nail, the benefits of having the right size hammer can be measured in terms of doctor bills not paid or time not lost from work because of a smashed thumb.

We have offered several suggestions for what could be called the benefits of having information. One is social power, personal or political. The

second is better decision-making. The third is monetary gain, personal earnings. The fourth is lives saved. The fifth is time saved. A sixth would be money saved. Are they at all related, or are we dealing with different kinds of information providing different kinds of benefits for different situations and different people? We shall work on this together.

(1) Merewitz, L.; Sosnick, S.H., 1971. The Budget's New Clothes. Chicago, Markham Publishing Co., 1971.

(2) Dürr, W.T., 1976. "Information and Power: Information as a Key in the Political Process". Baltimore Region Institutional Studies Center, University of Baltimare, Maryland, 1976.

STATEWIDE PLANNING FOR LIBRARY DEVELOPMENT -- THE INFORMATION GAP

Ronald Linehan, Planner
Texas State Library
Austin, Texas

BACKGROUND

Statewide planning for library development actually got started in the 1930's in the United States. In 1936, 45 states had plans on file for statewide library development. In 1940, a National plan was developed under the auspices of The American Library Association. The plans were all very ambitious ones whose proposals sound ironically contemporary. The four outstanding characteristics of both the national plan and the state plans were: (1) strong state financing of libraries, (2) county and regional consortia of libraries, (3) multitype library cooperation and (4) statewide training and certification for librarians.

The plans were notable in several respects. They were developed primarily by professional associations, they were products of a one-time master plan process and they were not implemented largely because of Adolf Hitler.

The present situation is very different. Hitler is no longer around, and planning for library development is generally an ongoing function of state agencies.

Long range planning became an ongoing function for statewide library development with federal legislation in 1969. This ongoing function was also heavily imbued with evaluation-ie, measurement of progress toward stated goals.

This led to some attempts to "model" planning...to assure that all the states would develop comparable plans. In 1971 the Ohio State University Evaluation Center instructed state library or agency directors in a process for planning referred to as CIPP (context, input, process and product). The bulky ambiguity of the process in the literature attests to its inapplicability.

There is, at the present time, a renewed interest and emphasis on planning information services not only at the state level but the national level as well with the efforts of the National Commission on Libraries and Information Science spearheading the efforts.

THE GAP

My concern is that in the provision of information services we have little in the way of information with which to plan -- we have an information gap.

Let me put this in perspective by giving a definition of planning, and to do that one needs to first talk about social policy or public policy development. Public policy development is generally referred to as the process of setting social goals, determining how goals will be achieved and how to measure success. Planning, I think, is an adjunct to that process. Planning is directed at two things: facilitating the political process, ie, getting people together at the right time and providing the participants with pertinent information so they make wise decisions on goals and programs.

The pertinent information should obviously correspond to the issues regarding decisions to be made. Generally there are two kinds of issues which can be addressed in planning for social services. The first issue is one of definition of service, the second is the issue of delivery system. The first involves a question of the priority of a service, which may be necessitated by changing societal needs, replacement of a service with alternative technology or the displacement of need. The second issue concerns not the basic definition of a service, rather the method by which a service is delivered. Both issues usually exist simultaneously... one may be more pressing. However, the information needed to address these issues is distinct.

To exemplify the two kinds of issues, the chart on the following page illustrates some decision options which may be considered by a state agency charged with planning and evaluating statewide library development.

The program areas listed are obviously not exhaustive; they are examples of program areas among which the professional community must select and fix a funding level, or reject. The kinds of information needed to assess the importance of such a program area or to assess the success of a program in a specific area are generally similar in each case. However, in most cases, the only data available falls under the category of current resources --number of volumes, number of films, etc. The key information dealing with the resource/use correlation and projected use is simply not available at the state level.

While benefit/cost analysis may be applied to assist in helping decision-makers evaluate choices among delivery systems, and this kind of information is becoming more available in the literature, choices among program areas is largely a matter of intuition.

The dearth of evaluative mechanisms concerning broad program areas of library development statewide results in a limited ability to plan for the future. Traditionally, evaluation is usually

PROGRAM AREA	INFORM. NEEDED	DELIVERY SYSTEM OPTION	INFORM. NEEDED TO ASSESS DELIVERY SYSTEM
To increase # of books available in state	Current resources intensity of use resource/use relationship Deficiencies-current intensity of use projected	Direct grants for purchase of materials Reimbursed reciprocal lending	Unit costs Political feasibility Other constraints
To increase # of films available in state	"	"	"
To increase building capacity	"	New state legislation	
To increase professional staff in state	"	Direct grants for personnel System grants for substitute professionals Grants to universities for OJT projects	"

TABLE 1

accomplished on the basis of data collected for four types of analysis:

1. Comparative Analysis
2. Extrapolation
3. Consumer Preference Analysis
4. "Functional" Analysis (correlation analysis, etc.)

For the last decade, comparative analysis of two sorts has been the primary method utilized in statewide planning. Growth has often been measured as regards ALA and state library association standards, or vis a vis some normative statistics developed in the state. There are exceptions. Hawaii's plan (1968) relied heavily on extrapolated resource data and New York's 1969 study incorporated functional analysis with extrapolated population data as well.

While a wide number of new resource and output measures have been posited, such as Deprospo, Beasley and Altman's library profile measures and Morris Hamburg's exposure hour measure, there needs to be more concentrated and realistic work done to make these measures useful for statewide planning and evaluation. There needs to be a shift from the comparative sorts of studies to more "functional" kinds of analysis to provide decision-makers with a tool for estimating the results of their funding decisions. As accountability to legislatures becomes ever more a pressing issue, the need to show how program allocations affect use and users will increase.

SMERC:

AN ON-DEMAND EDUCATIONAL INFORMATION CENTER

Susan Long
San Mateo Educational Resources Center
Redwood City, California

How much would it be worth to you, a classroom teacher, school librarian, school board member, or building principal, to have access to an on-demand, professional, rapid, and up-to-date educational information center--a center which could keep you informed about trends in education, furnish curriculum materials for use with your class, provide background literature on current school problems? For the classroom teacher, SMERC might furnish guides for teaching the metric system; for the school librarian, a list of high interest, low vocabulary books; for the school board member, examples of budgets from similar districts; and for the principal, an interpretation of the most current legislation on keeping student records. SMERC can quickly locate and supply resources for all of these requests. Let's look briefly at the history and philosophy of the center, its users, its policies and procedures.

Unable to obtain federal funding for an information center because the initial proposal was too grandiose and the scope of the second request was too narrow, the San Mateo County Office of Education decided in 1969 to independently fund such a center for the teachers and administrators within its jurisdiction. Since its inception, the San Mateo Educational Resources Center, SMERC, has expanded to serve educators in twenty-three other California counties, five states, and several education-related organizations such as the California Teachers' Association central staff and the Industry-Education Council of California.

SMERC is a synthesis of several types of educational information systems which have emerged since the late 1960's. Like the retrieval systems, we inform our users of the existence of documents relating to their specific concerns. Like the dissemination systems, we assist educational agencies in sharing their published materials by including them in our resource banks and making them available to all educators. SMERC fits most closely, however, the characteristics of educational information systems identified by Brickley and Trohoski in The Evaluation of Educational Information Centers(1). These are: (1) reliance on a major data base which is then supplemented by locally acquired materials; (2) formalized interface procedures with clientele, usually through linkage agents or some process of negotiating inquiries; and (3) a user oriented product. At SMERC, we rely on the ERIC system as our primary resource and supplement it with our "fugitive file." We have linking or SMERC agents within each of our contracting agencies who are trained at SMERC in the negotiation process; and we recognize that our function is to supply the kinds of information needed by our clientele, not just information which we deem appropriate.

But we've gone one step further and it is that additional step which makes our center unique. We subscribe to the idea that an information system, to be valuable to educators, must provide one-stop service. That is, our researchers locate the materials related to the client's request, and then provide the actual documents which best fit the client's stated need, making any further information seeking unnecessary. This document accessibility is a key factor in our success as an information center. The value of accessibility was documented in the 1976

study of information users conducted by Paul Hood at the Far West Laboratory(2). Teachers, principals, and local and state educational agency administrators rated this accessibility as the most important feature of their primary source of information.

Access to the SMERC system is through a contract for services with the San Mateo County Office of Education. The standard agreement for services entitles the contracting agency to: (1) unlimited search services for all educators within the agency's jurisdiction; (2) a specified number of copies of all SMERC publications; (3) training seminars at the center for the person identified as the linking agent by the contracting agency; and (4) three field presentations to the agency's teachers and administrators explaining the services and how to use them. The system allows for flexibility. For example, personnel in one county office might make heavy use of SMERC services, while personnel in another might use it sparingly; one county office might send several people to the Fall and Spring SMERC training sessions and another might send none; one county office might use field presentations as an inservice program for their teachers or county office staff, and another might handle the explanation of the services by themselves.

Despite the above "options" as to the amount of use by an educational agency, the cost does not vary. The amount paid to subscribe to our services is currently based on average daily attendance and is now set at fourteen cents per student for full services for one year. Greater use of the services provided by the terms of the contract is the way to reduce the per request costs. For example, a district of 25,000 ADA would pay $3,500 per year whether they submitted one hundred search requests or 1,000. Ignoring for the moment all benefits except search services, this means that the cost per search for one hundred searches would be $35.00 per search and, for 1,000 requests, these same searches would cost the agency just $3.50 each. Another kind of cost versus benefit was brought to our attention from a high school in a member school district. The school paid $315 as its share of the county office's contract costs with SMERC last year. But one search submitted by the principal resulted in a saving of $3,000 on insurance costs for that school alone. Needless to say, that principal feels that he owes SMERC several years of membership even if he should never use the service again, which is unlikely after his success.

Within the SMERC network, agencies handle the contract costs in three ways. First, a county office of education or a state department of education may pay the entire cost for all educators in all school districts within its jurisdiction, providing it as a service from that agency to their educators. Those using this system often provide their own search forms and letterhead for use by our researchers, thus assuring that their office receives credit for the service provided. In many cases, teachers and administrators are unaware that the research was actually done at the San Mateo County Office of Education in Redwood City, California. Second, an agency may pro-rate the cost to member schools based upon ADA so that all eligible for the services share the costs. In this system, the agency signing the agreement for services collects from the individual districts or schools and is responsible for paying SMERC. In the third method, charges to participating schools are higher than actual SMERC charges and the extra fee supports the local information center which acts as an intermediary unit sometimes adding local information, and providing field presentations regarding the services provided by both centers.

Once money matters are settled, the users begin submitting requests. The

two heaviest user groups are classroom teachers and principals or other building-level administrators. Though there is a disparity in the size of the groups, limited research at our center shows that equal numbers of search requests are submitted. Even in the state contracts which make the services available to both state education agency personnel and local practitioners, the majority of the requests come from teachers and principals and not from the department of education. The previously mentioned Hood Study lists preparing reports, articles, or speeches as the purpose for seeking information by both of these groups. It is SMERC's experience, however, that the requests for assistance with an immediate problem far outweigh research requests which might be used as background for a paper or speech. The teacher who just learned that several EMR students were to be mainstreamed into her class seeks our help in locating materials on mainstreaming in order to assist these exceptional children to adjust to the regular classroom, and to manage her time to accommodate both groups of children. The principal facing an upcoming parent conference regarding the new open-record legislation checks with us to be sure he has the most recent analysis of the law. This is not to say that a professional paper or speech may not result from the information ultimately, especially if information gleaned results in a successful attack on the problem. But paper writing and speech making do not appear to be the primary purpose for seeking the information.

The clients we serve play a most important role in our center. They initiate the search process, define a specific problem or need, identify any limiting features, and set guidelines for the researcher. We recognize that each person coming to us for information has his or her own reason for making the request, and we, therefore, respond individually to each request, attempting to provide the documents which accurately and adequately answer the client's question. The individualized response package, tailor-made for the client, is another unique feature of the SMERC service. The teacher looking for information on setting up a woodshop in her third grade self-contained classroom would receive different materials from a middle school teacher looking for guidelines for setting up a woodshop in his school. The class structure, age of the students and even budgets would be considered by the SMERC researcher if the teacher mentions it as a limiting factor.

In order to make access to educational materials quick and easy and to encourage educators to rely on our information services, a system of linking or SMERC agents has been established. Within each contracting agency at least one person is assigned responsibility for accepting, negotiating or clarifying, and sending the requests to our research staff and for returning the completed search package to the client with any necessary explanation. Thus, the person serves as the link between SMERC and the user.

Unlike the linking agent described in the literature on information systems, the SMERC linking agent does not play the field or change agent role. That is, SMERC linking agents do not attempt to assess the needs of the clientele, but rely on the client to recognize and identify his educational problem and approach the linking agent for information. The linking agent, in turn, helps the client to develop a concise statement of his problem for the SMERC researcher.

The linking agent participates in an intensive, two-day training seminar at SMERC in order to gain familiarity with our services, resources, policies, procedures and personnel. The linking agent meets and works with the SMERC research analyst assigned to the contract and our analyst remains the contact person when the linking agent returns home. The chain with the client linked to the research analyst has been highly successful for us and we believe this personalized service is an

important factor in the high usage of our services.

The impact of an individualized response and the ease of submitting a search request is reflected in our volume of search requests. SMERC clients generate more than 18,000 requests for search services annually, and that figure is growing. The average search response package contains ten microfiche, twenty pages of photocopy and a computer printout listing fifty document abstracts. Annually, then, we send out in excess of 180,000 microfiche and 360,000 pages of photocopy, all focused on education-related topics!

To meet this demand, the center has a staff of twenty-four permanent employees. The nine research assistants are organized into three search units coordinated by research analysts. Each unit is assigned responsibility for specific contracts. The researchers and analysts have varied educational backgrounds with the majority being either teachers or librarians. No attempt is made to create subject specialists. Rather, search specialists are trained to find information in response to any education-related request. The search units work in conjunction with a reproduction group. This unit, as the name implies, is responsible for duplication of all microfiche and photocopied documents sent in response to client requests. In addition, they maintain the microfiche collection, file back all original microfiche used by the research staff, several thousand a day, and transform newly acquired resources into microfiche format. As you might suspect, the clerical support staff is also an integral part of the office operation, assisting each search unit by dating each incoming search, mailing out all response packages and keeping the office records regarding use of the service.

Because the primary function of our center is to provide search services, the resources housed at SMERC are vitally important. We draw upon these to respond to an estimated 100 requests per day. SMERC's primary resource is the Educational Resources Information Center, ERIC, the national resource bank of educational materials. Containing over 140,000 documents in microfiche format and an approximately equal number of journal citations it is searched for all requests.

To supplement ERIC materails, we have created our own data base of fugitive materials, educational information which is not found in any other known collection. Known as FIDO, the Fugitive Information Data Organizer, this collection of 5,500 documents focuses on locally-produced educational literature. Originally, FIDO had a decidedly California orientation; but as the center has expanded, so has FIDO's coverage. For the teacher, SMERC stores collections of classroom-oriented materials such as California Curriculum Guides, Kettering Foundation Unipacs and Santa Clara County Learning Centers. To round out the microfiche and journal collections, the professional library contains approximately 25,000 volumes, including classic as well as contemporary educational literature. In response to each search request--the school nurse looking for nutrition learning activities, the superintendent interested in education finance reform, or the teacher designing a new report card--we draw upon well over 250,000 documents in our effort to find the right materials for each person.

Searching a resource bank of this size requires computer assistance. The center's research staff uses our in-house system, COSMOS, which stands for County of San Mateo On-line System. COSMOS is an interactive computer system which enables the researcher to create sets of resumes on a specific topic, review them on a display screen, and restructure the search stragegy until a satisfactory number of on-target citations are located. The interaction of the researcher and the computer system produces

an individualized, tailor-made computer printout for each client. For example, the researcher answering a request for curriculum guides for gifted elementary school students may begin by isolating all documents indexed under the term "gifted." There are 973 resumes. Next, the gifted resumes are limited to those identified as curriculum guides, leaving 183 resumes. Finally, those resumes which are also for elementary school students are extracted and 37 resumes remain. After reviewing these to make sure they are appropriate for the request, the researcher may print the resumes and include the printout in the search package along with copies of the most pertinent documents. Although the process is sophisticated, a computer search of this nature would take less than five minutes.

To encourage our clientele to use our services and to make it easier for them to do some of their own searching, SMERC produces several types of publications. Among these are NewsNotes and Resource Guides. NewsNotes are designed to be a quick look at innovative educational topics. Typically, the NewsNotes contain six or seven topics each edited by a SMERC researcher. A recent edition contained the following ideas: book reports; minicourses; hyperactivity; energy conservation; motivating underachievers and the role of the principal. Resource Guides are prepared as indepth examinations of one topic of major importance and interest to educators. This year we have published Resource Guides on Learning Disabilities, Declining Test Scores and Small Schools. Though the audience for each of these is somewhat limited, we have found that they are highly valued by our clientele and that school districts often rely on them as inservice education tools.

Now that I've told you how fantastic we are, perhaps it would be of value to briefly describle factors which have caused us some difficulty. One of the major problems we face is staffing which is controlled by the San Mateo County Civil Service Commission. SMERC must hire and fire based on these guidelines and educational research assistants may be hired only as permanent, full-time employees. Our workload, however, fluctuates greatly. Because of our school-based clientele, we receive a huge influx of searches in October after the school year gets underway. With only a brief decline during December because of the Christmas vacation, the demand remains high until March. Instead of a bell-shaped curve, our search demand curve looks much like the Golden Gate Bridge with two very tall towers. As a result, we are faced with either being over-staffed during the summer months or under-staffed during the peak periods. As a financial matter, we opt for being slightly under-staffed and this sometimes causes us to fall behind our promoted search turn around time during the periods of high demand.

Being largely self-supporting, SMERC also has difficulty freeing funds for planning and development. Charges to our clients are set to cover only our services because we feel this is the only way in which we can keep the costs within the means of most school districts. Consequently, little remains for research and development. We hope to find a solution to this problem in order to keep our services up-to-date and valuable.

Ideally, one-stop educational information centers like SMERC could be developed throughout the country. A network of centers supported by county, state or national funds would enable all to participate in shared research and development projects. The NIE capacity building grants to state education agencies to facilitate development of dissemination capabilities could be the first step in developing such a network. SMERC is now subcontracting to perform the search services for these projects, thus paving the way for a truly national network of sharing and cooperation. In the meantime, SMERC's success and growth over the last eight years is demonstrating the continuing need for a one-stop, on-demand information service for educators.

References

1. Brickley, Richard R.; Trohoski, Carolyn V. 1976. The Evaluation of Educational Information Centers. Bethesda, Md.: ERIC Document Reproduction Service: 1976 December (ED099424).

2. Hood, Paul D. 1976. Comprehensive Education Information System Requirements Study: Data From the Final Report to NIE. San Francisco: Far West Laboratory for Educational Research and Development: 1976 September.

INTRODUCING THE CONCEPTS OF ON-LINE BOOLEAN SEARCHING WITH A MINI-COMPUTER BASED SYSTEM

Gerald Lundeen
Graduate School of Library Studies
University of Hawaii
Honolulu, Hawaii 96822

ABSTRACT

A mini-computer based on-line bibliographic searching package, written in BASIC, was developed for providing library school students and practicing librarians with an introduction to the concepts of bibliographic data bases, Boolean logic, and on-line searching. The package consists of programs for file input and modification, generation of an inverted index, and a search program. The search program provides the essential features of commercially available systems though the number of options and the complexity of the protocol is reduced. Advantages of a mini-computer based system as an initial introduction to the essential concepts of on-line searching are discussed.

INTRODUCTION

On-line searching of bibliographic data bases is becoming a rather common practice in libraries. Once the domain of the special Sci/Tech library, now we have a growing number of data bases in the social sciences and humanities also, and the use of these services by many academic libraries as well as some public libraries.

It is important that library school curricula include instruction and training in this area. Some approaches to this have been described in recent literature. (1,2,3) This training might be offerred in specialized courses in information storage and retrieval (1,2) and/or in courses concerned with information sources in the various subject areas (bibliography of science and technology, social sciences, humanities, business and economics (3), etc.).

Actual experience in doing searches at a computer terminal is invaluable for this kind of learning so schools need access to realistic computer based systems for educational use. The two major providers of on-line bibliographic search services, Lockheed Information Systems and Systems Development Corporation, offer access to their systems at reduced rates for use by library schools for instructional purposes. This makes possible the provision of instruction and experience with these systems which some schools may not be able to afford at the regular commercial rates. We make use of these services in our program (in the advanced reference courses and in the information storage and

retrieval course) at the University of Hawaii. In addition we have developed a mini-computer based on-line search system for preliminary instruction in the basic concepts of on-line searching with Boolean logic. This paper describes the general features of this system and our experience with it.

THE SYSTEM

The set of programs making up the searching and file maintenance system were written in BASIC and are run on the Hewlett-Packard 2000/ Access mini-computer at the University of Hawaii. An effort was made to allow for flexibility in record structure. Within a single file record organization must be constant but from file to file the number, kind, and order of the fields making up a record may vary.

Searching can be via an inverted index or by string searching of specified fields. Indexed fields may vary from file to file. (The major file used in our system is indexed by authors and subject descriptors.) Search statements may consist of single terms, which will be searched for in the inverted index; or a string and field code for a string search; or, previous search statement results may be combined with one another or with a new search term by the use of the Boolean operators AND, OR, NOT (represented by the symbols *, +, -).

The system provides the capability to display neighboring terms in the inverted index and of browsing in the index by scanning up or down in the alphabetic listing.

Date of publication is also searchable -- before, during, or after a specified year.

Records satisfying a given search statement may be printed at the terminal with a variety of formats. Any combination of fields may be specified or if none is specified a default format (defined for that file) is used.

The system also provides the possibility of having the output sorted alphabetically on any of the fields in the record (for example the results could be printed alphabetically by author, or title, or in call number order).

The program allows 20 search statements per search and if that is not enough, there is a KEEP command to save and "restack" some statements so the search can continue beyond 20 questions.

For first time users the program provides on-line instructions on the use of the program.

In addition to the search program, the system includes a

number of file maintenance routines for input of new records, modifying existing records, and generating the inverted index. The first record of each bibliographic file is used as a file label. This contains information for the various programs about the record structure of the file, the names of associated files, what fields are to be indexed, what (if any) field contains the date, etc.

CONCLUSIONS

Various versions of the system have been used in the Introduction to Information Storage and Retrieval course at University of Hawaii G.S.L.S. The main bibliographic file with which the students interact is one made up of readings related to the course. There are currently approximately 150 entries in the file. During the course the students are each assigned a small number of citations for which they are to assign subject descriptors and a brief annotation. They then enter these into the file, using an interactive prompting program. Thus we have a growing data base containing items of interest and the students have the experience of working at both ends of the process.

Most of the students enjoy the experience and find it useful. Those that have had the chance to use one of the commercial systems subsequently find that the prior experience with the mini-computer system makes it easier to become comfortable with "the real thing."

Our experience indicates four main advantages to using a mini computer system:

1. Cost: the mini-computer approach is much less expensive even with the reduced rates from the commercial systems. Our present cost for the use of the Hewlett-Packard 2000/Access system at the University of Hawaii is $1.00 per hour, compared to $15.00 or more for educational rates from the commercial vendors.

2. Not only is the user charge important from the point of view of cost but is perhaps more important in terms of the student's state of mind. Knowing that he/she is consuming $15.00 or more per hour while learning to cope with a large and complex system is likely to create some anxiety and consequently a lower level of learning for the average student; whereas at the rate of $1.00 per hour in local use charges (this is in actuality a method of allocating computer time within the University rather than a real cash output) the student is less likely to worry about costs and will be able to concentrate on the concepts to be learned.

3. The mini-computer based system is less complex to interact with while still introducing the basic concepts

involved. At the beginning the typical student knows little if anything about computers, terminals, system protocols, or about Boolean logic, set theory, and search strategy. By keeping the system relatively simple, the student can more readily learn the basic concepts involved, after which the transition to the large systems will be essentially a matter of learning the protocol of these systems, the additional "bells and whistles," and the characteristics of the data bases made available.

4. Such a system allows the searching of a variety of data bases so that it could be used in several courses in the library school curriculum as well as with continuing education workshops and demonstrations for practicing librarians. Mini data bases representing several of the important data bases in a variety of subject areas could be constructed as well as locally generated data bases. Local data bases might include the in-house practice catalog collection to provide training in the structure and use of machine readable cataloging.

It might be noted that an on-line search system such as is described here could also be used in the context of personal indexing schemes, replacing card files, edge-notched card systems, etc. With the increasing availability of increasingly powerful microcomputers (in kit form or fully assembled) at rapidly decreasing costs, such an application is becoming more and more feasible.

REFERENCES

1. Davis, Charles H.; Gryczkowski, Carolyn L. "On-line Searching in the Classroom." In: Compendium of Presentations of the 5th ASIS Mid-Year Meeting, Nashville, TN, 1976. Washington, D.C., American Society for Information Science; 1976. 171-176.

2. Bourne, Charles P.; Anderson, Barbara E. 1976. Observations on the Use of the Lockheed Dialog System for Laboratory Work in a Fall 1975 Course on Computer-Based Reference Services at the UCB School of Librarianship. Berkeley, CA, University of California, Institute of Library Research, January, 1976. 11p.

3. Cveljo, Katherine. "On-line Retrospective Searching as a Method in Teaching Information Sources, Services and Research in Business, Finance and Management." In: Compendium of Presentations of the 5th ASIS Mid-Year Meeting, Nashville, TN, 1976. Washington, D.C., American Society for Information Science; 1976. 177-186.

EFFICIENCY OF COMPETING MODELS FOR PREDICTING BOOK CIRCULATION BY SUBJECT: A THEORETICAL AND EMPIRICAL BASIS FOR ALLOCATION

William E. McGrath
University of Southwestern Louisiana
Lafayette, Louisiana 70504

ABSTRACT

It is shown that different sets of variables can be equally or nearly equally efficient in predicting book circulation by subject. Proportion of variance (R^2) for two basic models containing various combinations of 21 variables were tested. The first, called shelflist model, treated number of library books whose subjects match those of academic departments, as a control variable. As such it accounted for 67% of the variance, thus entering the equation first. With this model, four separate significant sets emerged each accounting for approximately the same amount of variance (10%). They were: (1) number of faculty, hard/soft, masters enrollments; (2) masters enrollments, majors upper level, hard/soft; (3) credit hours x total enrollments, hard/soft; (4) majors, upper level, total majors, enrollments, lower. In the second, shelflist was constrained by defining the dependent variable as the proportion of shelflist circulated. With this model, three separate significant sets emerged. These were: (5) masters enrollments and hard/soft (20% variance); (6) hard/soft and Ph. D. program (16%); (7) upper level majors, and credit hours x total majors (15%). In each of the tests, no other variables were significant. Of the 21 variables, 10 did not appear in any of the 7 sets. Any of the seven sets could be used in an allocation formula. The deciding criterion for choosing a set may depend on the convenience of collecting data for each of the variables in the set.

EFFICIENCY OF COMPETING MODELS FOR PREDICTING BOOK CIRCULATION BY SUBJECT: A THEORETICAL AND EMPIRICAL BASIS FOR ALLOCATION

William E. McGrath
University of Southwestern Louisiana
Lafayette, Louisiana 70504

BACKGROUND

An important problem in supplying information is how to determine which subjects to emphasize in building library collections. Widespread interest is partly practical—for allocating budgets, e.g.,--and partly theoretical. This interest is currently stimulating considerable research by librarians and nonlibrarians. Recent examples are Thomas Pierce (1), a recent Ph. D. in economics at the University of Notre Dame, Jacob Cohen (2), an economist at the University of Pittsburgh, and Mary Beilby, a Rutgers University doctoral student.

To determine which subjects to emphasize, librarians need first to know how to measure the use of their collections, and then to define the indicators of that use. One measure has been previously defined by this author (3,4): the number of books circulated in one semester whose subjects match those of the academic departments in the university.

If a good understanding of the indicators (variables) were attained, one would then have the beginning of a rational or theoretical basis for allocation. Pierce and others have noted that many earlier formulas merely listed what their authors assumed to be significant variables, without testing their significance. Both Pierce and the author have done extensive analysis of a number of variables in separate environments, with the work of one tending to corroborate the other.

Pierce has developed an allocation formula, employing the variables he found to be significantly correlated with use of books by academic departments in the Notre Dame University Library. These variables were departmental library size, total departmental majors and a departmental scale called the hard/soft index. This finding was similar to that of the author who found shelflist (departmental library size), number of masters level enrollments in each department, and the hard/soft index to be significant. Pierce's formula uses average book prices, regression coefficients for total majors and the hard/soft index, and departmental library size as a control.

PURPOSE

The purpose of this paper proceeds from the assumption that more than one set of significant predictor variables could be employed in such a

Table 1

Variables Showing r^2 for Two Approaches

	Variable	Shelflist Model[a] r^2		Proportion of Shelflist Circulated[b] r^2	
1	Number of faculty	0.145	*	0.058	ns
2	Hard/soft continuum	0.100	*	0.10	*
3	Pure/applied continuum	0.0001	ns	0.001	ns
4	Life/nonlife continuum	0.029	ns	0.029	ns
5	Number of credit hours being taught	0.088	*	0.029	ns
6	Number of enrollments, lower (freshmen, sophomore)	0.072	*	0.033	ns
7	Number of enrollments, upper (junior, senior)	0.084	*	0.033	ns
8	Number of enrollments, masters level	0.134	*	0.11	*
9	Number of enrollments, Ph. D. level	0.039	ns	0.028	ns
10	Shelflist	0.668	*	———	—
11	Ph. D. program (a dichotomous variable)	0.036	ns	0.027	ns
12	Masters program (a dichotomous variable)	0.009	ns	0.003	ns
13	Bachelor's program (a dichotomous variable)	0.015	ns	0.007	ns
14	Circulation (dependent variable)	———	—	———	—
15	Number of majors, lower level	0.061	ns	0.043	ns
16	Number of majors, upper level	0.119	*	0.076	*
17	Number of majors, masters level	0.020	ns	0.01	ns
18	Number of majors, Ph. D. level	0.026	ns	0.02	ns
19	Number enrollments (sum of (6-9)	0.104	*	0.04	ns
20	Enrollments-credit hours (product of 5 and 19)	0.121	*	0.04	ns
21	Total majors (sum of 15 - 18)	0.060	ns	0.037	ns
22	Major-credit hours (product of 5 and 21)	0.05	ns	0.023	ns

a. When variables are entered after shelflist
b. When variables are entered first, with proportion of Shelflist circulated as dependent variable
* Sig. at .95 level

formula. Whereas earlier writers often assumed that all possible variables must be employed for a complete formula, modern statistical techniques show that many variables in a large set may contain the same information—i.e., may be highly correlated. Thus to use them all would be redundant and probably invalid. One could then select whichever is convenient for other reasons--ease of data collecting, for example. In a theoretically sound application, significant variables would account for 100%, or nearly 100%, of the variance in circulation.

METHODOLOGY

Most or all of the variables in Table 1 have been mentioned by other writers as having or potentially having significance as determinants of collection use. To the author's knowledge, however, the 2nd, 3rd, and 4th were not systematically employed in the library literature prior to the work of this author, who gleaned them from studies by Biglan (5,6) and Kuhn (7). Departmental values for these variables appear in (4). Variables 20 and 22 are widely employed by state agencies in the funding of universities, and by federal agencies in the reporting of national statistics. They were introduced here to test the utility of the student-credit-hours concept in predicting library use.

In all cases, the unit of analysis is the academic department. It is important to note that borrowing patterns of individual faculty members or students are not analyzed.

Analytical method was stepwise regression, from which the principle statistic, proportion of variance (R^2) was determined. Coefficients appearing in the computed regression equations would be applied to allocation procedures, as in the Pierce formula.

In the author's previous studies, and in Pierce's, number of books circulated by subject was the dependent variable. Both treated shelflist, i.e.—the number of books owned by the library in the given subject areas—as an independent variable in its own right. As such, in both works it accounted for an overwhelming amount of variance, demonstrating that it must be incorporated or controlled in any subsequent analysis or application.

In this study, shelflist is controlled in two separate tests: (a) by allowing it to enter into the regression equation first (shelflist models) as in the previous analyses, and (b) treating it as a constraining variable. In the first, variance remaining after that of shelflist is the target of analysis. In the second, following the suggestion of Professor Jacob Cohen (8) of the University of Pittsburgh, the presence of books in the library is regarded as a necessary condition for their circulation. The shelflist is therefore treated as a constraining variable, much the same as the econometrician does when he treats demand for money as a homogeneous function of money income. This is accomplished simply by defining the dependent variable as the proportion of shelflist

circulated—i.e., circulation divided by shelflist. Variance of circulation is then freed from that part of the shelflist not circulated.

RESULTS

Table 2 lists significant sets of variables for the two approaches. General sequence in each succeeding trial was to eliminate one or all significant variables from the previous trial. Thus, variables 1, 2, and 8, significant in the first trial for the shelflist model, were eliminated insubsequent trials. Note that, in multiple regression, variables not significant by themselves sometime become significant after other variables are entered.

In the shelflist model (Table 2a), number of faculty, hard/soft and masters enrollments were significant to the exclusion of other variables, accounting for 10% of variance after shelflist. When variable 1 was deliberately suppressed, variables 8, 16 and 2 were significant. The only other significant set to emerge was that containing variables 16, 21 and 6. There is very little difference in the total amount of variance that each of the four sets account for, even though each set contains a different combination of variables. None of the remaining variables (Table 1) were able to contribuite any additional significant variance, nor could they be recombined in any significant way.

Table 2b shows three significant sets, each containing only 2 variables. The one with the largest variance (20%) is that containing masters enrollments and the hard/soft characteristic. The other two sets, containing different combinations of variables, are about equal. As in Table 2a, no other combinations were significant.

DISCUSSION

The principle difference between the two approaches is that the best shelflist models require two or three variables before exhausting variance in each set. Models for proportion of shelflist circulated require no more than two, and account for twice as much variance. It would seem therefore that the latter would be preferable in an allocation formula.

Given the choice of sets with equal proportions of R^2 the deciding criterion may rest upon the convenience of the significant variables. Data for enrollments and majors, for example, are relatively easy to collect, since they are universally collected.

The hard/soft concept, on the other hand, is tedious to measure initially. If done well, however, it should be fairly stable from year to year and should not require remeasurement for some time.

Data for Ph. D. program is easy to collect. A department either has a Ph. D. program, or it does not. Utility of this variable should be viewed with extreme caution, however, for at least three reasons: (1) it is not significant taken by itself (Table 1); (2) when it is significant

Table 2. Relative Efficiency of Significant Models

Table 2a. Circulation as Dependent Variable

Trial	Variables Tried		Significant Models	Cumulative R^2		Variable Convenience[a]
1	all	1	no. of faculty	0.048		A
		2	hard/soft	0.079		D
		8	masters enrollments	0.103	Total	B
2	all but 1	8	masters enrollments	0.045		B
		16	Majors upper level	0.079		B
		2	Hard/soft	0.103	Total	D
3	all but 1, 8	20	Credit Hours x Total enrollments	0.040		C
		2	Hard/soft	0.082	Total	D
4	all but 1, 2, 8, 20	16	majors, upper level	0.0396		B
		21	Total majors	0.0599		B
		6	enrollments, lower	0.0921	Total	B

Table 2b. Proportion of Shelflist Circulated as Dependent Variable

5	all	8	masters enrollments	0.11		B
		2	hard/soft	0.20	Total	D
6	all but 8	2	hard/soft	0.10		D
		11	Ph. D. program	0.16	Total	A
7	all but 2 and 8	16	upper level majors	0.076		B
		22	credit hours x total majors	0.15	Total	C

a. A = easy, B = relatively easy,
C = somewhat tedious D = tedious

—i.e., when hard/soft is deliberately suppressed—it still accounts for a low amount of variance (less than 7%); (3) as a dichotomous variable (1 if a Ph. D. program, 0 if not) it contains little information—i.e., low variability; if all departments in a university have Ph. D. programs, they would all have the same score, 1, and thus the variable would be meaningless.

Variables consisting of the product of credit hours and enrollments (or majors) are somewhat tedious because they require precomputation, and would thus be somewhat messy to include in a pragmatic formula. It should also be noted, that data for circulation and shelflist are somewhat tedious to collect.

Notable is the lack of significance for several variables which many writers previously asserted ought to be included in an allocation formula (table 1). Also notable is that the maximum amount of variance accounted for in the constrained model was only 20%, a far cry from the desired 100%. This finding indicates that there is much we do not know about library circulation. A more converted search for meaningful variables is highly desirable.

REFERENCES

1. PIERCE, THOMAS. The Economics of Library Acquisitions; A Book Allocation Model for University Libraries. Ph. D. Dissertation, University of Notre Dame, 1976.

2. COHEN, JACOB. Book Cost and Book Use, The Economics of a University Library. Paper presented at the Pittsburgh Conference on Resource Sharing in Libraries, Sept. 29-Oct. 1, 1976, University of Pittsburgh.

3. MCGRATH, WILLIAM. Predicting Book Circulation By Subject in a University Library. Accepted for publication in Collection Management, 1977.

4. ———. Relationship Between Hard/Soft, Pure Applied, and Life/Nonlife Disciplines and Subject Book Use in a University Library. To be published, 1977.

5. BIGLAN, ANTHONY. Characteristics of Subject Matter in Different Academic Areas. Journal of Applied Psychology. 58:195-203, 1973.

6. ———. Relationships between Subject Matter Characteristics and Structure and Output of University Departments. Journal of Applied Psychology. 57:204-213, 1973.

7. KUHN, THOMAS. Structure of Scienfific Revolutions. Chicago, Ill. University of Chicago Press, 1962. 210 p.

8. COHEN, JACOB Personal Correspondence, Jan. 3, 1977.

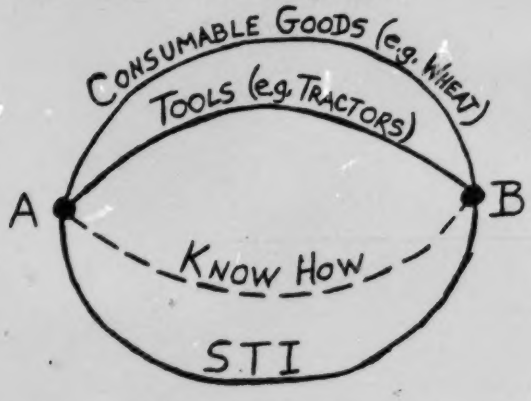
WORLD TRADE IN INFORMATION AND KNOW-HOW

Prof. Peter P. M. Meincke, Physics Department
University of Toronto, Toronto, Ontario

I would like to examine the concept of the value of information in the context of several factors which are playing an increasingly important role in decision making in countries around the world. Some of these factors are leading to a world in which the transfer of information trade between countries will be more in information and less in actual goods. If the future brings such a globe of almost self-sufficient states rather than a world order in which countries specialize in production of wheat and cars or oranges and electronics and supply the rest of the world's needs, it will be essential to develop a mechanism for putting a value on information.

The first part of this paper will be devoted to an examination of the factors which appear to be moving us in the direction of trade in information rather than goods. The second part will outline the reasons why it is so essential to develop an adequate pricing and control mechanisms, and explore some possible solutions in the light of what is successful today.

Let us first describe with the aid of a diagram the ways in which countries interact. Figure 1 shows two countries A and B exchanging goods and information in four very broad categories.



At one end of the spectrum, the exchange between countries A and B is in the form of raw materials or directly consumable goods, like food and clothing. In the aid context wheat or milk would be shipped from one country to the other. In the trade context some balance is sought with one country selling an indigenous product like wheat while the other sells oil. Prices are established in a variety of ways ranging from free market supply and demand to price fixing by cartels.

The next broad category is the transfer of basic tools, for example tractors or manufacturing technology that can be used to produce goods in the first category. In the aid context, this category was considered to be superior to the first because it increased the capacity of the developing country to meet its own needs for food and clothing, increased employment, and reduced the day-to-day dependency on the supplier. The developed countries were also happier because, initially at least it increased export markets and stimulated employment in manufacturing sectors. Again the prices or values of the goods transferred are well-known, and understood even if payment is not actually made to the developed country. In the trade context, one country finds it can manufacture T.V. sets and cameras more cheaply and sells to other countries. Another specializes in cars. There is of course the opportunity to balance payments with a mixture of category 1 and category 2 goods, but many countries that export raw materials soon begin to wonder why they shouldn't process and manufacture as well.

But there is growing dissatisfaction with exchange in both these categories. The trend toward self-sufficiency is very strong especially in the field of energy. The U.S. has mounted a major program to eliminate its dependence on foreign energy sources. A recent document, "Energy Policy Review" submitted to the National Economic Development Council by the Energy Secretary of Great Britain expects that Britain will become self-sufficient in energy by 1980.

The trend to self-sufficiency is very noticeable in other areas as well, especially in countries that have somewhat volatile political relationships with other countries. India's most recent five-year plan strongly emphasizes the goal of self-sufficiency in fertilizers as well as energy. Government policies clearly discourage the importing of goods and encourage local production even of the most sophisticated equipment.

Another factor is the rapidly growing interest in the concept of appropriate technology. Although much work remains to be done in refining the definition of Appropriate Technology, a number of characteristics are emerging. One characteristic is the need for the technology to be matched in quantity and quality to local needs and conditions. It should also make maximum use of local resources.

There are many examples of inappropriate technologies especially where some technology that appeared to be successful in a developed country was transferred without modification to assist a developing country. There are trucks and tractors rusting in fields because there is

no infrastructure to supply spare parts or maintenance. Another often-quoted example is the large-scale plastic sandal factory which succeeded in totally disrupting the local economy by throwing all the sandal makers out of work and decreasing the demand for leather, and creating hardship for the farmers. Moreover, the factory could only work at partial capacity because there wasn't a large enough market to absorb the production and the lack of maintenance infrastructure meant it was down much of the time.

A more advanced technological example is the modern telephone exchange installed in Delhi. Apparently it was imported from a country in which the average telephone call is just a few minutes long. The situation in Delhi is quite different because the telephone has become a major medium of social communication and the telephone calls last much longer so that the exchange simply cannot handle the situation.

It is an interesting question as to whether a nuclear power plant is more appropriate than a solar energy technology which could provide power for irrigation and to a widely dispersed village population.

More and more people are starting to think in terms of a more decentralized technology, a technology that has been designed with local needs and resources in mind. One can see this concept emerging both in high technologies such as interactive computing and pocket calculators and in the very successful work of the Intermediate Technology Development Group in London, England, which has designed a variety of agricultural tools and decentralized manufacturing processes such as a small glass-making plant.

There are many other examples which illustrate the growing trend toward a more decentralized technology that takes on the characteristics of the local society and is matched to local needs and conditions.

What happens to interaction between countries if the goals are self-sufficiency, making rather than buying and developing appropriate technology rather than using technologies from the developed world. Obviously such trends would greatly diminish the links through categories 1 and 2. The question is whether or not valuable interaction can take place in other ways.

The fourth category at the other end of the spectrum in this model is the transfer of scientific and technology information, the kind that appears in the literature and so is freely available. In principle such information is free although people are certainly used to paying for the vehicle for carrying such information through such techniques as journal subscriptions. It is the general philosophical tradition that such information is free that leads to the difficulty people have in realizing that other kinds of information have value and can have a price attached. (Spies of course have capitalized on this fact throughout history.)

The problem is that such information is not enough to do something with. Every country knows $E = mc^2$ but certainly can't make an atom bomb, at least so far. This suggests another category of interaction somewhere between the level of basic STI and tools. It is often referred to as know-how and it is the category that may well become most important. During a recent trip to India, I constantly heard about the need to exchange "know-how", so I would like to label the third category between Tools and STI "know-how".

There are a number of important aspects to the exchange of know-how. First there is

the need for an infrastructure,

a capacity on the part of the receiving country to take the know-how, introduce any modifications necessary to make the technology appropriate, and

translate it into real product. I believe it is this kind of capacity that organizations such as the International Development Research Corp. in Ottawa have in mind when they fund development projects.

A second aspect is that the resulting technology has a much greater chance of being appropriate because local needs and conditions can be taken into account, as the know-how is translated into product. This theme constantly emerged in the discussion, with Indian Government Officials, research workers, and private businessmen.

The most important aspect for the purposes of this paper is the cost of know-how. It certainly costs money to develop know-how, far more than it cost to discover the basic knowledge from which the know-how evolved. It is estimated that the ratio of development to basic research costs is in the order of ten to one.

In a very real sense, know-how is part of a company's capital investment just as valuable as the plant and equipment used to manufacture goods in categories 1 and 2. The return on this capital investment is through the sale of the tangible products.

If, as seems to be the case, the trend is to move from category 1 and even category 2, a very serious question arises as to how a return can be obtained for a capital investment in know-how. We do not expect companies to give away part of their plant. Can we expect them to give away know-how? Another difficulty is that there will be little incentive to invest in the development and transfer of new idea if there is no return on "know-how."

Therefore we are led to the conclusion that know-how must be assigned a value if the trends away from categories 1 and 2 are to be accommodated within the established economic framework. But the task of establishing this value is an enormously difficult one. One of the fundamental problems arises because in this context know-how becomes a product as well as a capital good and the cost of reproduction is negligible.

It is instructive to examine the ways in which know-how is transferred now and how prices are established in activities such as licensing, consulting, education and training. Most of these techniques involve transfer of funds, that is a value is placed on the information transferred (fees, tuition, etc.). For example, there are many interesting features about the licensing arrangements not the least of which is cultural imperialism. The most important aspect is that so far know-how seems to be transferred best through human to human interaction such as training programs with intensive hands-on experience. Anyone who has tried to build an experimental apparatus entirely from the description in a journal article or even a detailed thesis will know what I mean.

Can the value (price) of consulting services, education and training be reasonably well established through the normal market system? This appears to be true to some extent in the Universities but as more and more of the cost is borne directly by Government subsidy, the value (price) of a university education becomes harder to establish. Courses such as management seminars seem to be able to establish a value (price) on know-how but it bears little relationship to real marginal cost.

In view of the difficulties of putting an intrinsic value (price) on know-how and the fact that so far the most effective transfer mechanism seems to be through human contact rather than through packaged information, I would suggest that we build on the established practices of consulting and training to establish value (price) for know-how. It may even be that in the long run, when infrastructures and capacities are built up within each country, the majority of interactions could take place at level 4, the STI level where really only the cost of the actual transfer mechanisms need by recovered.

In the meantime, it is essential to develop a widely accepted mechanism for putting a value on know-how if the trends toward self-sufficiency and appropriate technology continue as strongly as they appear to be now.

A CORPORATE LIBRARY'S EXPERIENCE WITH

SELLING ON-LINE INFORMATION RETRIEVAL

Betty Miller
Information Specialist
David Mindel
Micrographics Supervisor
Calspan Corporation
P.O. Box 235
Buffalo, N.Y. 14221

Calspan Corporation, formerly Cornell Aeronautical Laboratory, is an independent research, development, and engineering company. As our former name indicates, our interests were originally in the aeronautical field, but, as time passed, we added other interests such as transportation safety, electronics, computer science, energy programs, human factors studies, etc.

Since Calspan's Technical Library must rapidly serve the information needs of these diverse fields, we became involved with on-line information retrieval in January 1973. At that time we began accessing three data bases of Lockheed's DIALOG. Gradually other systems were added and we now search approximately 58 files from six on-line services.

Based on the internal enthusiasm about on-line retrieval, the suggestion was made in the Spring of 1974 that perhaps we had a saleable commodity in our expertise with on-line systems. By this time, we had been searching for eighteen months and were averaging about thirty hours per month on-line.

Our first attempt at marketing was twofold:

1. We developed a brochure which was sent out in a bulk mailing, but this approach was not very successful.

2. We made a 16 mm movie which explained the advantages of on-line retrieval and briefly described the files we could access. This movie was originally shown at a demonstration-cocktail party given for members of the Western New York Library Resources Council. This method of advertising was only moderately successful. Later the movie was converted into an 8 mm cassette for showing on a Technicolor Showcase and in this format was displayed at several meetings.

For about nine months we continued performing commercial searches but did very little in the way of actual marketing. We had a few regular clients and occasionally would do a job for someone else. In March of 1976, we decided to launch an extensive marketing campaign. Our approach included 4 techniques:

A) Advertisements in library newsletters and local directories.

B) Individual demonstrations (using the portable terminal) to key people in area industries (i.e., technical directories, executives, librarians).

C) Slide presentations with an actual on-line demonstration when feasible for large groups such as technical associations, universities, and companies.

D) Exhibit booths at area malls and industry exibitions.

The clients obtained from these efforts and from other sources are shown in Table 1.

Note the large number of clients obtained through referral by our employees or ex-employees. In fact two of our good clients are ex-employees. They used the service frequently when they were at Calspan and continue to do so now on a commercial basis. Many clients also are referred to us by professional contacts, local librarians or others in our field whom we have met at various meetings.

The problem that we faced with regard to advertising was which vehicle to use. We selected library newsletters because we felt that the library community would be familiar with on-line retrieval. Another consideration was cost. Library newsletters are far less expensive than industry trade journals. The four clients received directly from the ads in the newsletter fully paid the costs involved. As for trade journals, we still are giving strong consideration to advertising in these and will probably do so in the near future.

We thought that individual demonstrations offered the best vehicle for marketing efforts. These presented an opportunity to bring the terminal into an office and actually interact with the potential client, leaving him with a full understanding of COINS' (Calspan On-Line Information Service) capabilities. The initial obstable to overcome in this approach is getting into the office. Approximately 50 phone calls were made to selected research and development corporations and businesses to schedule an actual demonstration. Twelve of the 50 opened their doors. The remaining 38 either had access to computer information retrieval already, were interested but not at this time, felt it didn't suit their needs, or were just plain not interested. A word of caution about telephone solicitation: try to talk to key people (technical directors, vice-presidents, etc.) otherwise you may spend a good hour fascinating a clerk, secretary, etc., and never get a demonstration.

The large group slide presentations sometimes coupled with a physical demonstration seem to have been the most profitable of all our efforts. Two relatively large accounts have been established as a result of such presentations. The success or failure of this kind of activity is also contingent upon your initial telephone contact. Should your contact be aggressive enough to promote the presentation to the corporate staff, helf the battle is won. However, if it is made with a person with little skill in conjuring up enthusiasum, don't count on a good turnout for your presentation.

We found that universities and colleges were not as promising as we thought. Generally speaking, those in need of information in an educational environment, are not under the deadline gun as those in industry frequently are. Consequently, while the faculty may be fascinated by the concept of on-line information retrieval, the cost of such a service will not be justified. A graduate student may be considerably slower in accessing information for the instructor, but he is also considerably cheaper.

In reference to exhibit booths our experience has been negative. COINS was exhibited twice; once at a shopping mall in the heart of a downtown business district and once at an industry show. No significant business was generated from either. After some analysis we felt that an exhibit might be worth while at a library orientated conference, i.e., SLA, ASIS, etc. We hope to display COINS at some such future event and enjoy a more fruitful result.

For Fiscal Year '75, the first year we were selling our services, we averaged 4.8 searches per month. This increased to 6.6 for Fiscal '76 and to 8.6 for the 10-month period March through December '76 since we have been marketing COINS more vigorously. [Figure 1] It is difficult at this stage to know how much of a market exists in our area. We have done searches for people as far away as Atlanta and Quebec, but all but a few of our clients are in the Western or Upstate New York area. One client accounts for 37% of our business and eight clients for an additional 41%. The remaining 31 clients provide less than 1% of business each.

Two of our earliest clients have since begun accessing DIALOG or ORBIT themselves and our best customer will probably do so in the near future. This is to be expected and brings up the question of why any library bothers to use our services. Some of them ask, "Can't we sign up with Lockheed or SDC ourselves?" Our answer is that of course they can but unless they are going to do a significant amount of searching every month, it will not pay them to buy or lease a terminal, train someone to use several systems, spend the time necessary to keep up with new data bases, new features, etc.

Because of our background, our service is oriented but not limited to science and technology and, therefore, most of our clients are in these fields. However we have done some work in business and the social sciences. [Our clients include high school students working on an NSF grant, a market researcher, a community college administrator, several businessmen, and others.] About 64% of our business is done with other libraries. These are mainly small technical libraries that do not feel that they can become involved with on-line searching themselves at the moment but still want to have the advantages of on-line searching available. Another 21% of our business is with small companies, mainly R&D, which do not have libraries of their own or have extremely limited ones. The remaining 15% includes students, teachers, lawyers, etc.

We stress our capabilities with on-line searching, but we will do manual searching and consulting if desired.

We also provide duplicate microfiche copies of NTIS and ERDA documents and have just begun offering document delivery service. With the exception of NTIS or ERDA reports or journals that we have in-house, we are subcontracting the document delivery work.

The question of confidentiality is vital to all information brokers. It is doubly important to us. We not only have to consider the possibility that two of our clients may be competitors but the possibility that one of them may be a competitor of Calspan. So there have been occasions when we have done a search for a commercial client and a very similar one for an in-house user within a few days of each other.

Currently we are in the process of revising our brochure to make it more user oriented. Our original brochure was geared to the information professional. This one will be slanted toward the individual researcher, businessman, etc. We will send these new brochures to any contacts we have made in the past and to any new prospects.

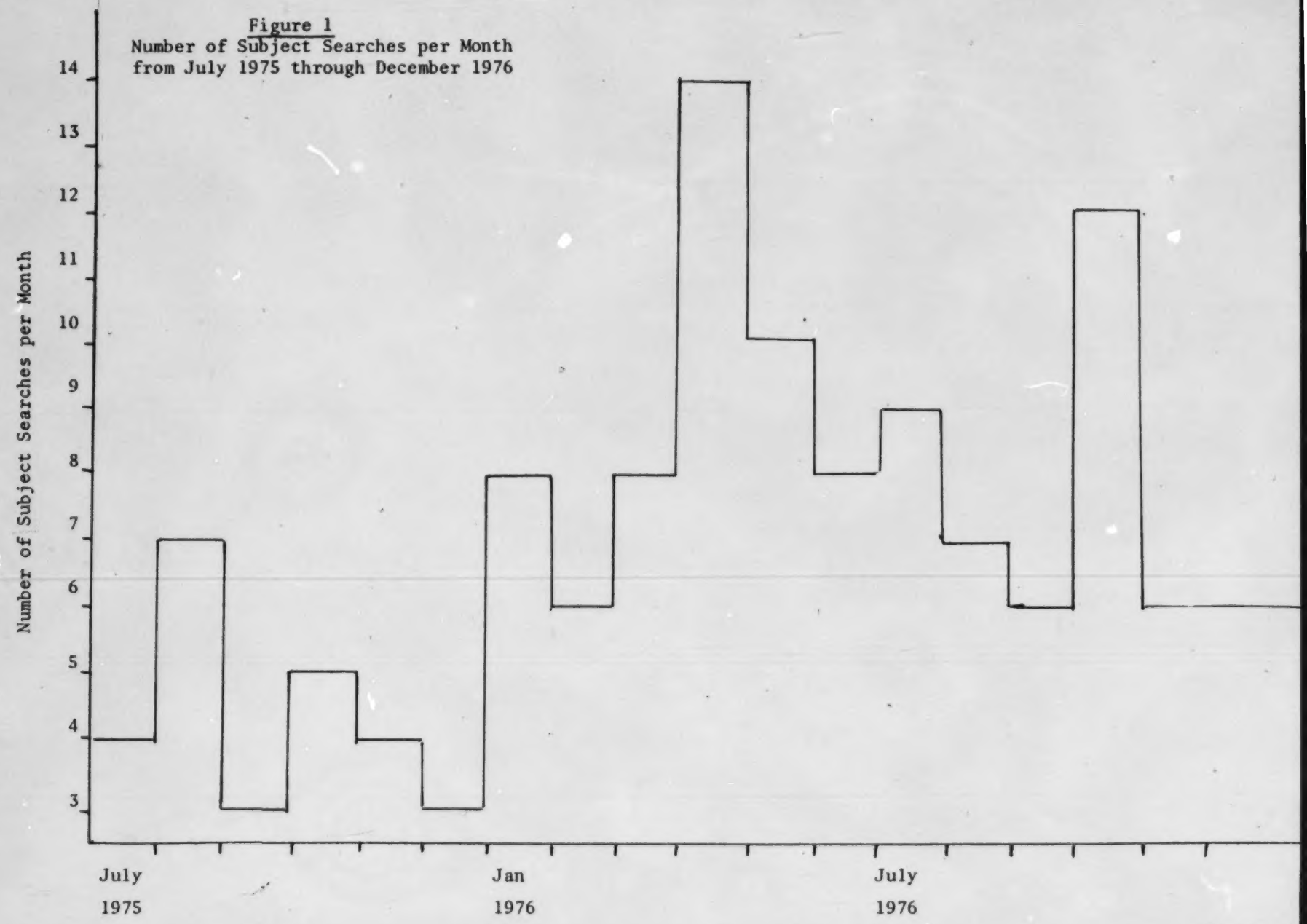
On-line information retrieval is still in its infant stage. As more and more people become aware of it and its capabilities, the larger and more identifiable becomes the market. Response may not be immediate but the seed has been sown and with the proper cultivation it will bloom in a fruitful manner.

Formal presentations/demonstrations	10
Referral by employees or ex-employees	9
Referral by professional contacts	8
Cocktail party/demonstrations	5
Special Library Association Newletter	4
General mailing	2
Listing in ERIC catalog	1
Listing in telephone directory	1

Table 1

THE NUMBER OF CLIENTS OBTAINED THROUGH VARIOUS CHANNELS

Figure 1
Number of Subject Searches per Month
from July 1975 through December 1976



PERIODIC MANAGEMENT REVIEW OF INFORMATION SERVICES

K. Musselman and J. Talavage
School of Industrial Engineering
Purdue University
West Lafayette, Indiana 47907

Descriptors: Information Services, Quality, Value, Effectiveness, Review Procedure

Managers of information services strive to provide their users with the most benefit that the service budget allows. Continuation of satisfactory service in an environment where the user needs changeover time implies the usefulness of a periodic evaluation of the offered services. A method is described in the paper which can assist the manager in such an evaluation. Each service is considered in terms of three attributes, namely quality, value, and effectiveness. For each attribute, several factors are measured by either questionnaire or are obtained from available data. These measurements are combined using established statistical or pattern recognition techniques. The result is a quality-value-effectiveness profile over time for each service. Such a profile may be used for problem analysis or for projection of future service usage and revenues.

INS AND OUTS OF AN IN-HOUSE EVALUATION - PROBLEMS AND SOLUTIONS FROM A WORKING INFORMATION SYSTEM

Joan Nugent and Deborah Carrow
The General Electric Company
Washington, D.C. 20024

Abstract

In-house evaluations, while necessary for effective management of information systems, are often set aside due to cost and time considerations. It is possible, however, to construct such evaluations using existing personnel, operating procedures, and resources. The experience of the National Criminal Justice Reference Service, a LEAA-funded international criminal justice clearinghouse with a data base of over 26,000 documents, is described as a case example of such an in-house effort. NCJRS recently undertook an assessment of its indexing and retrieval capability. The NCJRS evaluation approach is described, with emphasis on the practical solutions developed for such problems as devising a plan, availability of personnel, plan implementation, time limitations, and compounded workload strains. The NCJRS in-house evaluation is shown to be a practical and effective method of obtaining basic system performance figures.

INTRODUCTION

The National Criminal Justice Reference Service (NCJRS), is operated by the General Electric Company under contract to the Law Enforcement Assistance Administration (LEAA), to provide information on documents on all aspects of criminal justice: police, courts, and corrections. An extremely diverse population is served by NCJRS, with users ranging from criminal justice professionals and the academic community to the public at large. The NCJRS data base now contains 26,000 items, with the current acquisitions ranging up to 1,000 documents per month. As each document enters the system, it is abstracted, and assigned appropriate index terms from the NCJRS thesaurus, which contains some 2,000 descriptive and geographic terms. Each month the reference service receives over 1,000 requests for information by telephone, letter, and personal visit. Requests are as diverse as the user population.

NCJRS has the capacity to search its data base by use of index terms, by free text search of titles, annotations, and abstracts (TANAB), or by a combination of index and TANAB searches. When a request is received, it is forwarded to the information specialist whose expertise in the area of either police, courts, or corrections allows him or her to best address the question. After considering the nature of the request and interpreting

the needs of the user, the information specialist conducts a thorough search of the data base. The user is then sent document abstracts and bibliographic information tailored to his request.

Since the capacity of NCJRS to serve the criminal justice community hinges on its ability to index and retrieve documentation, NCJRS and LEAA agreed that an evaluation of this function would be conducted, and this provision was included in the NCJRS contract. Under the agreement, NCJRS was allowed a period of six months for submission of a plan, completion of an evaluation, and a report of the final results.

In determining an approach to the evaluation, we weighed the advantages and disadvantages of consultant-operated and in-house efforts, and concluded that NCJRS should conduct its own evaluation. Although the in-house approach would require the education of NCJRS staff in the field of information system evaluation, it was felt that the expertise gained would constitute an organizational resource that could be of value even after completion of the study.

LITERATURE SEARCH

A thorough literature search in information system evaluation techniques was thought to be the best starting point for our in-house assessment. Over 75 documents were studied in detail; considerably more were briefly reviewed. This time investment was deemed necessary, as it formed the basis for the entire evaluation effort. During this period the two person staff made frequent progress reports designed to inform management of the State-of-the-Art so that realistic goals could be set.

The most crucial problem faced during the literature review was a critical lack of documentation in the "how-to" aspect of information system evaluation. We had expected to find a great deal of material on these practical considerations, and found instead a large selection of references dealing with theoretical concerns, non-operational programs, or evaluations of systems vastly different from our own. A few references did prove to be invaluable, and much of our planning was based on these: Lancaster's Information Retrieval Systems, Lancaster's reports on the MEDLARS evaluation, and King and Bryant's Evaluation of Information Services and Products. We made an effort to obtain actual copies of all articles, reports, and books related to our subject, thus constructing a basic reference collection.

At the conclusion of the review, NCJRS staff and management had a workable knowledge of the possible evaluation approaches and the benefits and drawbacks of each approach. As we matched our objectives with the features of our system, it was possible to select those approaches most suited to our situation.

THE WRITTEN PLAN

A formal written plan was required under our contract; however, we found that committing the plan to writing served more than just fulfillment of the contractual obligations. This process forced us to establish our objectives, to examine carefully each available method, to clarify our reasons for choosing each procedure, to examine available resources, and to consider the impact of the evaluation on our system. The written plan was circulated several times among NCJRS staff and management for comment and criticism. This served the dual purpose of educating the NCJRS staff about the assessment process and alerting the evaluation staff to foreseeable problems and impacts. By the use of in-house resources and expertise, we were able to tailor the evaluation effort to the needs and operations of our system.

EVALUATION CONCEPTS

In its final form, the evaluation design contained the following basic elements:

1. The NCJRS evaluation would utilize actual operating conditions of the system; that is, actual subject requests, existing data base items, and existing indexing and retrieval procedures. Data would not be generated from artificial test conditions which might falsely represent actual performance.

2. The retrieval performance of NCJRS would be judged by the well-known measures of precision and recall. To calculate recall measures, a sample of data base items relevant to each request (the recall base) would be found without the use of indexing aids.

3. Document relevance would be determined by a panel of judges drawn from the criminal justice community.

4. The main thrust of the evaluation would be on measuring retrieval performance using index term searches. In addition, the performance of index and TANAB (free text) searches would be compared in a portion of the test cases.

EVALUATION PROCEDURES

These design concepts were translated into several specific operational elements. The test sample size was set at 92, a figure reached by the following method. After consideration of our evaluation objectives, we determined that a certainty factor of 90 percent and a permissible range of 10 percent would be adequate. Utilizing results of other system evaluations, we estimated that the coefficient of variation for the population to be sampled would be approximately .57, and calculated that a sample size of 92 would be sufficient. This figure was to be modified as results became available and as the actual coefficient of variation could be more precisely determined.

NCJRS developed a selection matrix to ensure that the test sample would reflect actual operating conditions. The matrix was constructed using statistical data reflecting incoming requests during the period of October 1975 to October 1976. The matrix (figure 1) included two basic request variables: mode of request (telephone, letter, or personal visits), and general subject area of request (police, courts, or corrections).

Number of Test Requests	Letter	Telephone	Visit	Totals
Police	19	16	2	37
Courts	9	7	1	17
Corrections	18	18	2	38
Totals	46	41	5	92

Figure 1. Selection Matrix

During the testing period, all incoming requests were sent to an evaluation coordinator, who randomly selected a sample of requests based on the predetermined matrix. Test requests were chosen during the period of January 11, 1977 to February 23, 1977. Each selected request was given to the appropriate information specialist, who translated it into a written subject statement. The information specialist then formulated a search strategy using index terms only and searched the data base to retrieve relevant items.

Concurrently, the subject statement was given to a special staff of persons assigned to find recall base items. The staff used non-index reviews of the data base holdings to select items related to the request. Items retrieved by the computer search and items chosen for possible inclusion in the recall base were recorded, and a master list combining both sets of items was produced. In instances in which the computer searches retrieved over 35 items, a random selection of 35 was included in the master list. Hard copies of each of the items on the master list were obtained from the NCJRS library, and these documents were used as the basis for all relevance assessments.

To make the necessary relevance judgements, NCJRS employed three graduate level students of criminal justice. For each test request, a judge was given the written subject statement and actual copies of all documents on the master list. The judge examined each document to assess whether or not it was relevant to the subject statement. All decisions were recorded and submitted to the evaluation coordinator. Using this data, the coordinator calculated the precision and recall ratios for each test search.

Twenty-three requests were submitted to a special procedure in which information specialists conducted two searches for the same request: one using index terms only and one using TANAB only. In these test requests, precision and recall values were calculated for both searching methods and were compared as part of the retrieval effectiveness evaluation.

EVALUATION PROBLEMS

The development and implementation of the steps outlined above was not one smooth process. The original plan was rewritten several times; the pre-testing period exposed problems in the test design which had to be corrected; and in addition, the test itself revealed considerable gaps between the idealism of the plan and the reality of the situation.

In translating the plan from paper to real life, it became obvious that centralized management of the process would be necessary to ensure staff cooperation and orderly operation of the test. An evaluation coordinator was therefore appointed to manage both the pre-tests and the actual evaluation.

A number of potential problems became obvious during the two scheduled pre-test periods. For example, we had originally thought that the judges would review abstracts instead of documents. During the pre-test we realized that this would not produce the best possible relevance judgements, and modified our procedure so that judges would review hard copies of the documents. Another potential problem area was the number of items produced in each test search to be reviewed by the judges. We thought that judges could easily review a great number of abstracts, but when we decided that documents would be reviewed instead of abstracts, this became unfeasible. Searches of the NCJRS data base can produce thousands of items. It was therefore decided that on those test searches containing more than 35 items, judges would view a sample of 35 items drawn from the search results. In searches containing 35 items or less, all the retrieved documents were examined for relevance.

In retrospect, we found that problems experienced during the test itself stemmed from the tension between our efforts to maintain a fair and realistic evaluation and our efforts to keep up with our normal operations while the test was being conducted.

In our decision to require that the judges review hard copies of each document we foresaw a shelving backlog. The magnitude of the problem became apparent only when we began the operation, which entailed for each search: locating the document on the shelves, delivering them to the judges, sorting them on return, and reshelving them. Over 5,000 documents were involved in this process during an eight week period. Our intent was to minimize disruption of ordinary service by allowing our holdings to be out of circulation for the shortest possible time. In order to retain control of the documents, we developed a stringent and time-consuming record-keeping procedure. We also hired some temporary shelving help, instituted preliminary sorting procedures, and finally held a "shelving party" attended by the entire fifteen-person document processing staff.

An evaluation invariably causes workload strains on all involved. This was particularly evident in the case of our information specialists who, in each of the 92 test searches, had to formulate and conduct at least two separate searches. A normal NCJRS response was sent immediately to the user, and a second search using index terms only was conducted for the test. In an additional 23 cases, information specialists were asked to prepare a response using only free text searches. This situation illustrates a typical problem resulting from an evaluation effort. Since no satisfactory solution was found, information specialists were forced to acquire working backlogs to be dealt with after termination of the testing period.

The most significant departure from our original plan lay in the method used to establish recall bases for each test search. Our original procedure called for manually searching our title listing for relevant items. We realized, however, that establishing a recall base by this method could prejudice the results in our favor, and decided that a more objective selection procedure would be to use our system's free text search capacity to form recall bases for the index term searches. Since the free text search does not scan the index terms assigned to each document, index terms were never viewed when selecting items for the recall bases. Four abstracters were chosen to perform the recall base searches. Although this method was more time consuming, the increased objectivity of this procedure was considered to be worth the increased investment of time and personnel.

EVALUATION BENEFITS

Was it worth it? Seven months after beginning the evaluation process, we would still have to say yes. We obtained valuable measures of our retrieval effectiveness: for example, the average index recall ratio was 64.4%; average index precision was 88.6%; TANAB recall averaged 41.2%; and average TANAB precision was 82.9%. Since comprehensive statistics were collected on each individual search, we also have a substantial body of yet unutilized data which can be applied in further analyses of our system.

In addition to attaining our original objective of measuring indexing and retrieval effectiveness, the evaluation produced a number of practical benefits. Management and staff gained a quality of insight into the nature of NCJRS operational problems and advantages which could never have been achieved by a non-participatory consultant-managed assessment. For example, the need for more communication between indexers and information specialists was illustrated during the formulation of recall bases, when these two groups had trouble defining their respective perceptions of the same term. Having problems highlighted in this manner was non-threatening due to the in-house nature of the evaluation. No one was coming in and finding fault with us -- we were discovering and rectifying our problems on our own initiative. This factor certainly helped to ease the usual resistance to evaluation experienced by many organizations.

The need for communication highlighted during the testing period was to some extent answered by the evaluation itself. During the planning stage, different functional groups within the organization were required to interact and to provide planning input; during the actual evaluation this communication process was increased. This benefit might never have been realized if an outside evaluator had handled the planning and coordination of the process.

A further result of our experience was the knowledge of information system evaluation gained by NCJRS personnel. This expertise is a valuable organizational resource and is a direct result of our time and personnel investment.

To complement the knowledge of our staff, we have acquired a substantial reference collection on information science evaluation theory and practice. We expect to enlarge upon this basic resource and to apply it to future evaluation needs.

Finally, we have accumulated an extensive collection of data on our system operations and performance. As a potential resource, this information offers us an opportunity for continuing insight into our retrieval capabilities and a basis for planning future studies.

CONCLUSIONS

We recognize that the in-house evaluation approach was not an easy route. In fact, its greatest value was that we gained experience as we gained our results. As we review our efforts we can now identify a number of basic precepts which may aid others considering similar in-house assessments.

1. As soon as possible, establish the long range goals of the evaluation and the major objectives leading to that goal.

2. Obtain a commitment from management on the goals and objectives of the evaluation. Solutions to inevitable problems will be facilitated by management support.

3. As you do your literature search, get hard copies of all pertinent documents. It is not sufficient just to take notes on the material as you will need to refer to it often.

4. Make a written plan for the evaluation. As implementation problems arise, the plan can be modified. Make sure that these changes are fully approved and documented.

5. Realize at the outset that you cannot anticipate every problem and prepare to meet unexpected difficulties with the most workable compromise or solution.

6. Be prepared in particular for the backlogs which the evaluation will create in the normal workloads of your organization.

7. Conduct several practice runs of the evaluation. The time will be well spent. You will discover problems which need to be solved before your test can run smoothly, and you will familiarize your staff with the mechanics of the operation.

8. Make use of all available personnel, assigning them to new or unusual tasks as the need arises.

9. Consider the time the staff spends on the evaluation as an organizational investment which gives the staff valuable insight and expertise.

REFERENCES

1. KING, DONALD W. and BRYANT, EDWARD C. The Evaluation of Information Services and Products. Washington, D.C.: Information Resources Press, 1971.

2. LANCASTER, F.W. "Aftermath of an Evaluation." Journal of Documentation, v 27 (March 1971): 1-10.

3. LANCASTER, F.W. Evaluation of the MEDLARS Demand Search Service. Bethesda, National Library of Medicine, 1968.

4. LANCASTER, F.W. Information Retrieval Systems: Characteristics, Testing and Evaluation. New York: John Wiley and Sons, Inc., 1968.

5. LANCASTER, F.W. "MEDLARS: Report on the Evaluation of Its Operating Efficiency." American Documentation, v 20 (April 1969): 119-142.

INTEGRATING LIBRARY ACQUISITIONS WITH THE CURRICULUM

Richard Palmer
School of Library Science
Simmons College
Boston, Massachusetts

UNUSED LIBRARY ACQUISITIONS

Recent quantitative use studies by Richard Trueswell (University of Massachusetts, Amherst) and Ward Shaw (University of Denver) have found that about 40 per cent of items acquired by college and university libraries are not used by faculty or students. As these findings become more widely known, acquisitions librarians will doubtless be seeking ways to ascertain in advance of placing orders those items likely to be used. Unfortunately, most librarians are not now well situated to find out which items will actually be used by faculty for instructional purposes or by students for learning needs. Nevertheless, as library acquisition budgets come under more intensive scrutiny, librarians will come under increasing pressure to reduce the number of unused items they acquire.

In an attempt to determine possible reasons for the high percentage of unused acquisitions in college and university libraries, I reviewed my own experience as a university acquisitions librarian. I recalled that I had never been consulted by any faculty members while they were designing curriculum or planning courses. My library colleagues and I usually learned of new courses from new issues of the university's catalog. Since we had no advance notice of new courses, or of modifications in courses, or of deletions of courses, we did not attempt to link our acquisitions policies or weeding activities to the curriculum--except in a very general way. We did not take the initiative to seek out active participation in curriculum planning at the dean or department head level. We did not seek to synchronize library acquisitions with curriculum on a course-by-course basis, though I believe now we should have. I also believe that had we participated in the planning and development of information resources for specific courses the library would have acquired a higher percentage of items that would be used.

LINKING ACQUISITIONS TO THE CURRICULUM

During the past two years I have been involved with experimental studies seeking to design and implement procedures linking library acquisitions directly to the curriculum. A team of information professionals, under my supervision, have been working with deans, heads

of departments, and individual faculty at a junior college with a very large student enrollment. Our goal has been to devise practical ways to integrate curriculum design and the acquisition process. The thrust has been to identify library items for which there is a known faculty or student demand, and buy these, rather than to buy items listed in publisher's promotional material or in conventional library selection aids. Instead of blindly acquiring items and wondering why they are not used, we endeavored to find out in advance which items would be used in specific courses by faculty and students and acquire them.

In our studies at the junior college we had the support of the administration, including the president, the academic deans, and the department heads. Following meetings with the president, we met with the deans and department heads and arranged meetings with faculty by departmental groups and individually. During these meetings we discovered that few faculty members had fully-developed syllabi for their courses. Even those that had such syllabi had seldom included specific plans for the use of library materials. The departmental chairmen, therefore, asked each faculty member to develop a complete syllabus for each course, listing goals and objectives, detailing each week's classroom activity and student assignments. When the syllabi had been prepared, the information professionals met with individual faculty members to discuss various instructional materials and equipment of potential usefulness. Few faculty members were aware of the great variety of materials that were available for acquisition by the library, and initially few expressed interest in having such materials acquired, even though the information professionals assured the faculty that they would make all necessary arrangements to acquire and deliver the materials and equipment as required.

The interaction between individual faculty members and the information professionals also included a comprehensive look at modes of instruction. The appropriateness of the mode to the content, the pacing of the course, and the variety of presentations were all examined. Efforts were made to achieve an interesting variety of modes, including case studies, films, slide shows, guest lecturers, tours, role-playing, simulations, tape recordings, videotapes, computer-based instruction, as well as other instructional modes. These consultations were held in advance of the school year so that materials chosen could be acquired before they would be needed. We found that faculty members were often passive, even somewhat reluctant in these endeavors. Had we not taken the initiative, no integration of library materials with the course syllabi would have been achieved.

LINKING ACQUISITIONS TO STUDENT ASSIGNMENTS

Since the team conducting the experiments recognized that faculty instruction is only half of student learning, the team discussed with

faculty members the need to provide for student access to library materials. Many of the junior college's courses were taught only in the evening when the library was closed or at sites many miles from the college library. The team discovered that faculty members were often unaware of the substantial variety of library materials which could be acquired and made available to students of varying abilities and interests--right in the classroom. The team assisted faculty in selecting for each week's student assignments various levels of library materials for students of different learning styles and capabilities. The library then acquired the selected materials. Throughout the courses students were introduced to self-teaching materials, films, slide collections, tape recordings, records, and computer-based programs, as well as other library materials acquired in these experiments.

The interaction between faculty members and the team is continuing. The information professionals continue to take the initiative to add other courses to the experiments. Although faculty members continue to be somewhat skeptical or reluctant at first, most of them report they have gained an increased awareness of teaching techniques and of library materials. We feel a continuing need to be aggressive and innovative. We need to market library materials and services. In the process we are learning what library materials to acquire--for a high percentage of use.

LINKING LIBRARY MATERIALS THROUGH BIBLIOGRAPHIC INSTRUCTION

Some faculty members involved in the experiments feared that the effect of our explicit selection of materials would straight-jacket their teaching and students' learning. To counteract any possible narrowing of the use of library materials, the faculty members and the team arranged for classroom instruction in library use and search techniques. These classroom presentations, conducted by the information professionals, avoided generalities, were focused on both new library acquisitions and already-held materials distinctly relevant to the course, and were timed so the sessions immediately preceded assignments requiring library use. Library tours followed the class presentations and the locations of particularly useful materials were indicated. The faculty members reported that students performed substantially better on their assignments following the team's bibliographic instruction.

INTERIM FINDINGS

We are continuing the experimental efforts in integrating library acquisitions and the curriculum because we believe interaction between the team and deans, departmental chairmen, faculty, and students is contributing to the following gains:

(a) The identification, in advance of acquisition by the library, of items faculty members plan to use or have students use in their courses.

(b) A reduction in the number of unused library acquisitions.

(c) A more cost-effective use of the library's acquisitions budget.

(d) A faculty more aware of library services.

(e) Students better able to make use of the library and library materials.

(f) Students achieving better test scores.

(g) Information professionals more directly involved in the educational process.

The experiments in integration of library acquisitions with the junior college's curriculum have produced benefits for all groups involved in them. We believe that similar efforts at integration can produce like benefits elsewhere.

EDUCATING MANAGEMENT STUDENTS IN INFORMATION UTILIZATION

Eugene J. Rathswohl
Department of Management
University of Texas
Austin, Texas

INTRODUCTION

There is a growing awareness in both the management and information science professions that management is an information intensive activity (1). With the growth of the information industry and the information arrangements and services made possible by such growth, there has been an increased tendency to view managerial effectiveness as being directly related to the ability to access and utilize information resources (2). On the other hand, there is behavioral evidence that managers as information users lack skills that would allow them to take greater advantage of the available information resources (3,4).

This paper describes the rationale and approach taken in designing two undergraduate management courses in the area of information utilization at the University of Texas. The courses are conceived of as user education courses and are offered as electives in the undergraduate management major curriculum. They are designed on the premise that undergraduate management students need to heighten their awareness of the possibilities provided by information resources and that as information users they need to develop practical skills for coping with the realities of existing information arrangements.

RATIONALE FOR COURSE DESIGN

A relatively simple conceptual framework has been used to think about course design. It consists of three major ideas: the information environment; the perceived information environment; and the interaction of the two. The information environment can be described as consisting of the various information resources (technologies, services, people) associated with the information generation and transfer activities of the economy. Such descriptions can include the growth of the information industry (5), the diversity and structure of existing information services (6), and the public policy issues related to the information industry (7). The perceived information environment refers to a specific user's orientation to the information environment. It is described by the user from the user's own point of reference. Each user develops over time a 'cognitive map' of the information environment which is integrally related to the developing epistemic style and autonomy of the individual (8). There are reciprocal relationships between the 'real' and the 'perceived' information environments that determine, in part, the ability of a user to effectively engage in information behavior. Information science has focused largely on improved design of the information environment. The linkage between improved information design and improved information use is, however, not automatic. It is not automatic because information use also depends on the active role the user must take in searching for and using information resources.

Two sets of assumptions are made. First, a user must take an active role in searching for and using information. The user must take initiative, for

example, in articulating information requirements, in deciding where to look for information, in deciding whether or not certain information is relevant. The concept of initiative suggests that the user ought to be able to operate at a certain level of autonomy; the user ought to be able to manage the information search procedures relevant to him. The concept of user initiative also suggests that information use is open-ended in the sense that an autonomous user ought to be in a position of discovering and inventing new information behaviors. The second set of assumptions relate to the information environment. The design objectives for information systems can be broadly stated as providing users the information they need, when they need it, and in forms that are readily useable. The users' operational requirements implied by this design philosophy is to interact with a continually growing, relatively sophisticated technological environment. On the other hand, the design complexity of the information environment may be preempting the user from taking the active role required. The complexity of the information environment may be precluding certain users from taking the initiative required of them to interact with information resources.

IMPLICATIONS FOR COURSE DESIGN

The rationale discussed above has had real impact on the design of information management courses in the U.T. Business School. Two undergraduate information management courses have been implemented during the past two years entitled Management of Information and Interactive Information Systems for Management. Unlike other business information system courses designed to educate students as information system designers, these courses have been designed explicitly for educating managment students as information users. Although learning to use business information is part and parcel of any business course, these two courses are the only courses in the management curriculum to address information utilization as a problem, per se. Both courses have focused primarily on learning to use library-based and computer-based information resources; the courses are being revised to also include the use of professionals, agencies, institutions, etc. as information resources.

The objective of these courses is not necessarily to support or reinforce the information use habits of management students not to modify their information behavior in some a priori way (9). Rather, the objective of these courses is to facilitate each user's initiative and autonomy in regards to information use. The rationale is that, given the complexity of information arrangements, a user's initiative and autonomy can be facilitated through improving the user's level of information awareness.

The problem has been in operationalizing the concept of information awareness. I have taken the approach of thinking of information awareness as a set of skills, i.e., a set of skills users need to develop to facilitate initiative. There is relatively little discussion in the information science literature about user skills beyond fairly general statements to the effect that users 'should be prepared and able to utilize man's cumulative store of knowledge through information technologies to optimize their problem solving skills' (9). From my perspective as an information scientist it is necessary to design these courses in such a way that the students actually develop skills through class activities, not just reading about them. It is necessary to have the students generate, from their own situation, a realization of the information skills they personally need to develop.

In light of the above considerations and other approaches to user

education (10,11), I developed a 3-step methodology for the courses. The first step is a presentation by an information professional to the students of a specific information resource(s). Examples of the resources discussed in class include: the U.T. Business/Economic library; the U.T. Government Documents and Microforms Collection; the ERIC, Psych Abstracts, INFORM, and Predicasts on-line data bases; OCLC; the New York Times Information Bank; and the Lockheed DIALOG system.

The second step of the methodology is for each student, subsequent to the presentation, to engage in an information search project utilizing that resource. The projects are done either on an individual or small group (3-4 students) basis. The objective of each project is to have students develop practical skills in using the resource. Examples of the type of projects assigned are the following:

- using the various published bibliographies, handbooks, encyclopaedias, periodical indexes, etc. in the Business/Economics library, a student compiles a bibliography of 10-20 items on the information needs of and information resources for a specific business area, e.g., business forecasting, corporate planning, mergers and acquisitions planning, etc.
- each student is asigned a specific business-oriented resource, e.g., Business Periodicals Index, Moody's Industrial Manual, Standard & Poor's Industry Surveys, Value Line Investment Survey, etc. The student writes up a 1-2 page report on the location of the resource, a short descriptive narrative of its content, the kinds of questions one could answer using the resource, and tips on what one needs to know to use the resource. Each student receives a copy of each othe other students' reports.
- using the various statistical publications in the U.S. Documents Collection, e.g., American Statistics Index, Statistical Abstracts of the United States, U.S. Census, a student compiles some appropriate statistics for a specific business-related issue in Texas, e.g., oil, gas, and goal reserves; employment trends in service industries; R&D expenditures for energy-related industries; family planning services; etc.
- each student selects a couple of recently published books and searches for them using OCLC. They experiment with all the options available for searching, e.g., author, author/title, ISBN, LCCN, etc.
- For both ERIC and Psych Abstracts data bases each student conducts an on-line search on a topic of his/her choice. These are non-delegated searches.
- For the New York Times Information Bank small groups conduct a search on a topic of their choice. This is a delegated search although they are able to observe the man-machine interaction.
- For the Lockheed DIALOG system one representative search is done on the Predicast data base for the class as a whole. This is a delegated search with a follow-up presentation in class of the man-machine interaction.

The third step of the methodology is for each student to write a 2-4 page narrative of the steps involved in using the resource(s) in the project. The objective of the write-up is to have each student explicitly identify the kinds of system design features, technological features, use procedures, intellectual factors, and possible problems and opportunities involved in using the resource(s) and how these factors effected the student's information

behavior. Subsequent to each project one class period is spent in class discussion about the information behavior factors involved in the project. It is through this discourse that students exchange ideas on the kinds of information skills they realized they needed to develop to complete the project. The discourse between students on their information behavior is a very important part of the methodology. It is through exchanging insights about information skills based on their own experience that students recognize how they might improve their own information behavior. Through the class discourse there emerges a general taxonomy of skills users need to develop that is based on their own interests as information users.

USER INFORMATION SKILLS

The following are the information skills identified by the students in the two information management classes in the past two terms. Two major classes of skills were identified: conceptual skills and praxis skills.

Conceptual Information Skills

Two conceptual information skills were identified: information style and reflexive thinking.

Information Style There is a certain consensus in information science that users ought to have some conceptual framework or model for defining one's information requirements and also of how to go about searching for and using information resources. Two models commonly used to structure a user's orientation are a research model and a library use model. The research model emphasizes information search as part of an overall inquiry strategy; the library use model emphasizes the availability of information resources. An underlying assumption of these models is that information behavior is a determinate process, that there are optimum, efficient ways of information search and use. This may well be an appropriate assumption to make in situations where problems and information requirements are a priori well-defined. However, many management situations are not a priori defined; rather, the manager must construct both the definition of the problem(s) and ways of coping with them. This suggests that having only one or relatively few conceptual frameworks within which to operationalize information behavior criteria may well preclude improved performance in other sectors. Although the research and library use models are good starting points, users also should be able to articulate information needs and search strategies of their own construction. The generation of conceptual frameworks for information behavior is the important issue here. Rather than imposing specific models on the user and thus risking impeding initiative, user education can focus on conditions that allow users to develop and test their own frameworks. The criteria of a user's information behavior effectiveness are appropriate only if they are defined from the user's perspective.

Reflexive Thinking A second conceptual skill is the ability to be aware of one's own information behavior. If users are responsible for developing their own information style they should be able to articulate the various aspects of the information environment and their perceived information environment that facilitate or impede their information behavior. Users must be conscious of information as problematic if they are to manage their own information behavior. Students found that they were not in the habit of thinking explicitly about their information behavior. They were not in the habit of thinking of themselves as information users. Part of the reason may

be the structure of business education. Although much course work is 'real-world' oriented in the form of case studies, business games, etc., the amount of thinking students must do concerning the information problems involved is minimal. All the pertinent information is given; what is hidden is the crucial process of being able to identify appropriate information in the first place. Students do not have to develop habits of identifying appropriate information, opinions about missing or inappropriate information, nor of the relative importance or value of information. Information is considered a given, not as a problem. Reflexive thinking shows how information is problematic and, thus, a management concern.

Praxis Information Skills

Two types of praxis skills were identified: intra-resource skills and inter-resource skills. Intra-resource skills are those skills required to search for information within a specified information resource, e.g., an index, a government agency, etc. Inter-resource skills are those skills required to search for information amongst a variety of information resources.

Intra-resource Skills

(a) Awareness of a resource as an information resource. This skill is being able to say whether certain types of information services exist or not. It is the ability to say what could be used as an information resource in a particular task or problem situation. Within a particular information milieu a user should be able to articulate what resources may be of pertinence. Possible information resources would not necessarily be limited to information technologies and services conventionally defined; they would also include people, places, and events. What is a pertinent information resource is defined by the user.

(b) Knowing how to get access to a specific information resource. This skill is knowing the general procedures for locating specific information resources. It is being able to identify: where or possibly where resources are geographically or institutionally located; access points to the specific resource; and general procedural requirements for gaining physical access.

(c) Ability to interact with information professionals. This may be the most important user skill, especially in the management field where information tasks are usually done through people, e.g, staff, vendors, and consultants. The ability of a user to articulate information requirements through discourse with an intermediary is critical. Users need to know what to expect of the intermediaries, what sort of help is available through them, and what intermediaries need to know to do their job more effectively. The more a user knows of the intermediary's task environment the better the dialogue between the two. Indeed, the user will be in a better position to evaluate the intermediary's performance.

(d) Hands-on skills. These are the commonly mentioned skills users need in non-delegated searches. They include such things as: finding, reading, and understanding instructions; translating their request into the language of the resource; and knowing how to use the resource to narrow or broaden their search. Each resource has its own intellectual and physical idiosyncracies; the user should be willing and able to identify and learn the skills needed to cope with those idiosyncracies. A user must be able to modify already learned hands-on skills to new or unfamiliar resources.

Inter-resource Skills

(a) Ability to go on to alternative information resources. Information resources are designed as interventions at relatively speicifc points in the information transfer process. Each information resource is necessarily bounded in terms of its coverage, types and forms of information contained, etc. A user should be able to recognize these bounds for what they are, i.e., technical bounds of that particular resource and not the bounds of the information environment as a whole. For example, after using a bibliographic index providing abstracts, a user should be aware that the full-text of the articles exist ahd how to go about gaining access to them. The link between these two resources is not obvious; one resource does not necessarily instruct the user in the possibilities of other resources. It is the responsibility of the user to link the use of the resources in ways useful to him.

(b) Ability to decide what information resource to search next. The ability to link the use of information resources depends on the user's intra-resource skills and how the user conceptualizes the role of information in the overall inquiry. The skill is the ability to hypothesize the form in which pertinent information may exist; it relies on the memory and experience of the user of past information search successes and failures. Pertinent information resources may be readily available; the issue is that the user must realize this in the real-time frame of his information need and in the praxis of the information search.

(c) Ability to decide when to terminate an information search. This skill is the ability to make a tradeoff between the relative costs invested in information search and the relative value of the information retrieved. This skill relates directly to the user's ability to operationalize and articulate his information criteria. The decision to continue or terminate a search should be just that, a decision. The user ought not give up out of being frustrated or ignorant of what might be done next to improve the search. Neither should the user terminate the search on the basis of premature success. This skill depends on how well the user conceptualizes the problem area in which information is being utilized.

REFERENCES

1. Harmon, G. 1976. "Information Science Education and Training." In: Annual Review of Information Science and Technology, Volume 11. Martha Williams, ed. Washington, D.C.: American Society for Information Science; 1976. 347-380.
2. Katz, R.L. 1974. "Skills of an Effective Administrator." Harvard Business Review. 1974 September-October; 52(5): 90-102.
3. Schewe, C.D. 1976. "The Management Information System User: An Exploratory Behavioral Analysis." Academy of Management Journal. 1976 December; 19(4): 577-590.
4. Lucas, H.C. 1975. "Performance and Use of an Information System." Management Science. 1975; 21: 908-919.
5. Dollotta, T.A.; et.al. 1976. Data Processing in 1980-1985: A Study of Potential Limitations to Progress. New York: John Wiley & Sons; 1976.
6. Parker, E.B. 1975. Social Implications of Computer/Telecommunication Systems. Stanford, CA: Stanford University, Center for Interdisciplinary Research, Program in Information Technology and Telecommunications; 1975 February. 62p. (Report No. 16).
7. Domestic Council on the Right of Privacy. 1976. National Information Policy: Report to the President of the United States. Washington, D.C.: National Commission on Libraries and Information Science; 1976. 233p.

8. Royce, J.R. 1975. "Epistemic Styles, Individuality, and World-View." In: Perspectives in Information Science. Anthony Debons and William J. Cameron, eds. Leyden, The Netherlands: Noordhoff International Publishing; 1975. 259-295.
9. Slamenka, V.; Zunde, P. 1972. "Science and Information: Some Implications for the Education of Scientists." In: International Conference on Training for Information Work. Proceedings. F.I.D. publ. 486. Georgette Lubbock, ed. Rome: Italian National Information Institute; 1972. 357-370.
10. Vernon, K.D.C. 1975. "Introducing Users to Sources of Information." Aslib Proceedings. 1975 November-December; 27(11-12): 468-473.
11. Klintoe, K. 1972. "User Profiles in Industry." In: International Conference on Training for Information Work. Proceedings. F.I.D. publ. 486. Georgette Lubbock, ed. Rome: Italian National Information Institute; 1972. 377-384.

INFORMATION AND DATA SYSTEM SIMULATION

A REPORT OF 1976 EXPERIENCES

Suzanne H. Roberts
General Electric Company
Gaithersburg, Maryland

If information were a commodity listed on the stock exchange, the market would indeed be bullish. Demand is increasingly rising, and to meet it, a growing number of technologies and systems are being developed, tested, and implemented.

Data processing systems are in the forefront of the information sciences field, and although they vary widely in scope and purpose, their end aim is the same: to acquire information, process it into a storage system, and provide for its retrieval. Implementation covers a spectrum ranging from a small, manually operated information distribution center to a large national information clearinghouse, a highly automated information storage and retrieval system, and a specialized data processor.

To make these systems as cost effective as possible, General Electric has developed a portable software tool that would assist in their initial design, operating management, and longer term adaptation. By combining experience in information center management and utilizing technologies of modeling and simulation, a simulator has been developed and is being used to help manage an existing information center. The simulator language, SIMTRAN, was developed by GE's Information Systems Programs from GASP IV, a general purpose FORTRAN-based simulation language. Provided are all the modeling and reporting capabilities available in GASP IV, as well as facilities for modeling process and resource management. Most significantly, the process (or transaction) oriented basis of SIMTRAN allows simulations to be quickly developed and modified from realworld descriptions of information systems characteristics without the necessity for large numbers of programmers.

SIMTRAN provides a preprocessor which allows the analyst to take a process oriented view and the system is envisioned as a collection of interacting processes, with the interactions controlled and coordinated by the occurrence of events. The advantages of this are:

- Taking a process-oriented view is a more natural way to express the structure and operation of an information system.
- Event definition and sequencing are accomplished by the simulator and are transparent to the user.
- System modeling is available at varying levels of detail, and the analyst can control variable values and ease of parametric explorations.

An analyst interested in making system level optimizations or tradeoff decisions can model an entire system with the relatively unimportant subsystems structured at a gross level of detail. The effort required to make future expansions or refinements to a model is also reduced. These capabilities were utilized in the development of the information center simulator used in the project. The simulator possesses inherent design factors which also aided the modeling activities.

They include:

- Generality - A generic simulator was desired so that it could be used to model a wide range of systems.
- Usability - The simulator is designed to be used by information specialists rather than sophisticated computer simulation users.
- Efficiency - Typically, many simulation runs may be made to analyze different resource tradeoffs or processing scenarios, and if a simulation run costs an excessive amount, the number of runs would be limited.

PERFORMANCE MEASURES FOR AN INFORMATION and DATA SYSTEM

Performance measures for an information and data system which were considered during the project can be grouped in four categories.

Responsiveness to the Customer/User - The response time between receipt of a request and delivery of the response, or receipt of data and its availability to a user/customer, is an important performance measure. The simulator provides the ability to determine what the response time will be with different staff mixes, under various workload conditions, using alternative equipment configurations, or operating with different priority schemes.

Throughput - The amount of processing or the number of jobs that can be completed given a certain system configuration and workload is provided by the simulator.

Reliability - The reliability of an information system is important. The effects of errors on subsequent operations is important to analyze, and the effectiveness of the system at identifying errors and correcting them would be valuable to model.

Quality of Information - While this may be the most important measure, it is also the most subjective and not presently within the capabilities of the simulator. However, a subjective measure can be established as to the data quality by analyzing the number of items that make up a response, the amount of time required to prepare the response, and the amount of reference material searched to provide the response. All of this information is available from the simulator.

Other information that is reported by the simulator and available for performing analyses includes such statistics as the queues at each resource, resource utilization, and transaction status.

MODEL DATA

The model used as input to the simulator consists of five different kinds of information: simulator run options, workload descriptions, resource descriptions, process descriptions, and variable descriptions. The workload descriptions specify the mix of transactions to which the information and data system is to be subjected. The resource descriptions specify the resources available in the information system. Both personnel and equipment types of resources were used in the model. The process descriptions describe the functions, interrelationships and quantitative behavior of the various processes which exist in the information system.

MODEL DEVELOPMENT EXAMPLE

The National Clearinghouse for Alcohol Information (NCALI) was chosen as the target system to be modeled. The basic functions performed in the NCALI operation are common to many clearinghouse operations and, in most cases, only minor modifications are required to model other information centers.

An information center is a dynamic system comprised of several interacting processes. The processes at NCALI are divided into the following major categories:

- Information Acquisition and Repository
- Information Analysis, Selection, and Abstracting
- Data Base Maintenance
- User Cultivation and Request Analysis
- Integrated Outreach Program development

A functional description of the National Clearinghouse for Alcohol Information (NCALI) is shown in Figure 1. This figure not only displays the functional components of the information center but also provides an indication of the scope and breadth of various products and services provided to the user community.

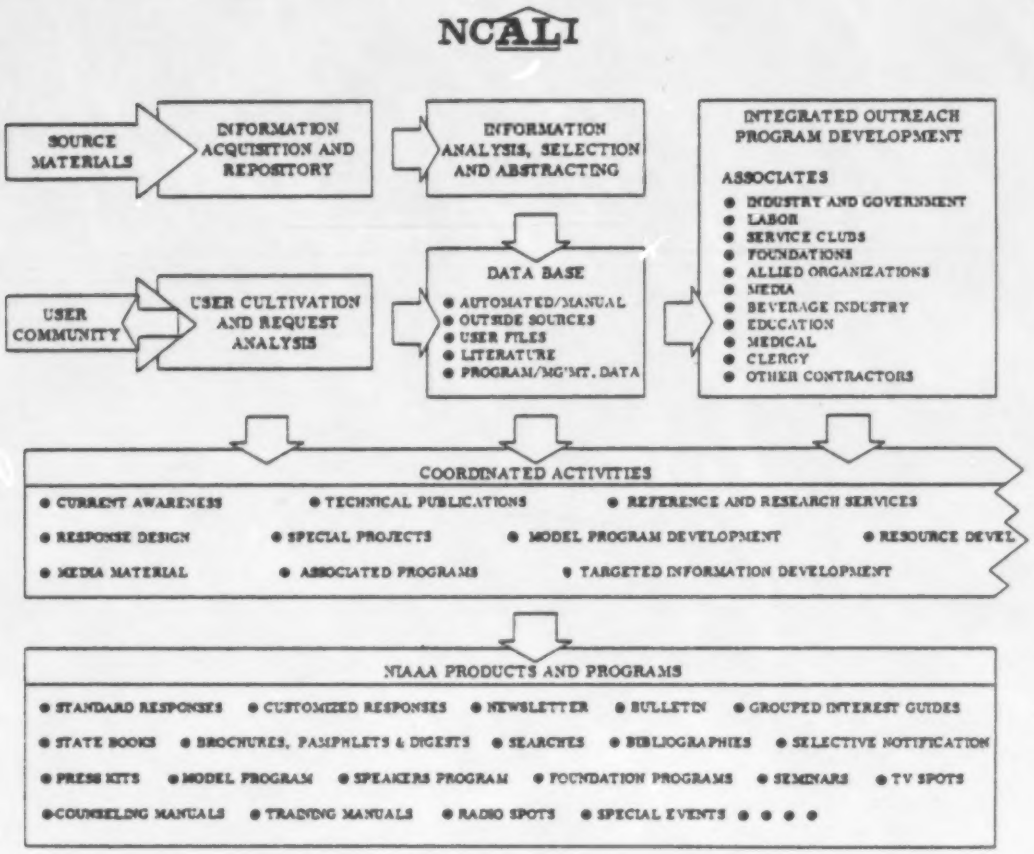


Figure 1. NCALI

DESCRIPTION OF THE INFORMATION CENTER MODEL

Through the analysis of the major functions being considered for modeling and the interviews with the operations personnel, the following processes were identified:

- Acquisitions and Document processing
- Request Processing
  - Standard Requests (those that can be filled with standard inventory items)
  - Special Requests (those that require additional analysis and reference activities)
  - Subscription Requests (those that add users to the regular publications and services of the information center)
- Product Development
- Selective Notification of Information
- Distribution Services
- Warehouse Services
- Computer Services

Workflow diagrams were developed for each of these functions which is represented by one or more process descriptions using the constructs provided by the simulator. The overall structure of the information center model is shown in Figure 2.

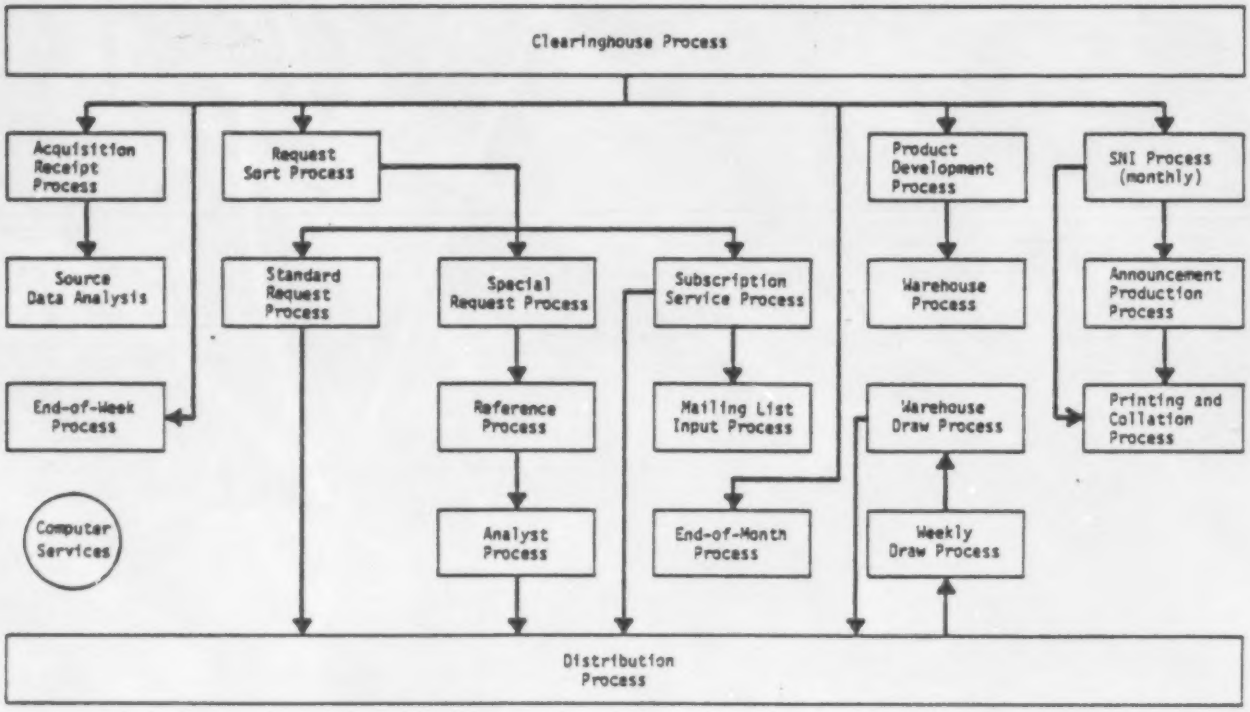


Figure 2. Overview of Information Center Model

In developing the information center model, a subset of the categories described above was considered. The core functions (information acquisition and repository, information analysis, selection and abstracting, data base maintenance, and request analysis) of an information center operation were modeled. Two other functions, user cultivation and integrated outreach program development, represent essentially independent groups in the clearinghouse and involve less definitive tasks such as user interaction and cultivation in the field, conference arrangements, etc. While important functions, they are less common to all information center operations, not as quantifiable in nature and, since they are independent activities, were not considered in the model.

MODEL DATA COLLECTION

A large portion of the effort in the project was involved in data collection, which was an ongoing activity throughout most of the project. The data was used in the development, fine tuning, and validation of the model. Initially, more detailed descriptions of the processes that made up the major functions of the information center operation had to be derived. A data collection form was developed and interviews with the operational personnel were conducted. Work flow diagrams were then developed for each process identified and discussed with the operational personnel in a second series of interviews for confirmation of their ac-

curacy. Data about processing times, number of items handled, costs, transaction types, demand frequencies, etc., were also collected at this time. The personnel were asked to describe one of the following in reference to activity performance time:

- A constant value
- A function of a variable (e.g., the number of items processed)
- A distribution (e.g., a normal distribution with a minimum of 5 minutes, a maximum of 25 minutes, a mean of 15 minutes, and a standard deviation of 4 minutes)

Once the functional areas of the information center were described by flow diagrams and data collected on the resources and times required to complete activities, the process descriptions were described in terms of the simulator input format. Twelve process constructs, shown in Figure 3, are available for describing the system. The constructs are used to manipulate resources, collect statistics, control other processes, and define the data flow within the process being described.

TYPE	FUNCTION
BEGIN	BEGIN A PROCESS
END	END A PROCESS
RQST	REQUEST A PROCESS
RLSE	RELEASE A RESOURCE
DELAY	DELAY FOR TIME TO PERFORM A TASK
CALL	CALL ANOTHER PROCESS
ENTRY	INFORMATION USED AT ENTRY OF ANOTHER PROCESS
RETURN	RETURN TO A PREVIOUS PROCESS
START	START A NEW PROCESS
ASSIGN	ASSIGN A VALUE TO ANOTHER VARIABLE
IF	BOOLEAN LOGIC OPERATOR
STAT	COLLECT STATISTICS

Figure 3. Process Box Types

Process descriptions were written by using the simulator constructs, resources were described, and variables defined and given appropriate values. In some cases it was more efficient to describe a functional area by more than one SIMTRAN process, depending on the batching and flow of data. The procedure of validating the model involved debugging and fine tuning the process descriptions so that they functioned according to expectations. This activity was accomplished by short duration simulation runs. All logic, coding and other errors were corrected after each run, and another simulation conducted. The workload to which the model was subjected represents the current demand on the information center. The request sort was initiated each day with variable quantities to represent the receipt of the standard, subscription, and special requests. The acquisition receipt process was initiated each day to model the acquisition of documents daily. The selective notification process was initiated monthly and the product development process weekly to begin the development of 10 products.

The end of week, end of month, and the weekly draw processes were also initiated to update the data files. Once it was determined that the total simulator and the model were operational, the simulation times were expanded to simulate a typical monthly operation, as shown in Figure 4. The output reporting capabilities of the simulator were used to provide statistics on the queue build-up at all activities, the utilization characteristics of each resource, and the processing rates of all products and services. The results obtained were compared against an established validation criteria and the historical data available from operational records. Results obtained from a one-month simulation are shown in Figure 5. To further validate the model, several experiments were conducted by varying different parameters of the model to determine their effects on the system's performance. Results obtained from these simulation runs were analyzed and compared against the expected NCALI performance under similar conditions. Operational personnel at the clearinghouse were consulted and verified the validity of the simulation results.

SIMULATION INPUTS

Run Time = 9600 Minutes

	MIN	MAX	RATE	DISTRIBUTION
Standard Requests	255	425	Daily	Uniform
Subscription Request	30	60	Daily	Uniform
Special Request	12	24	Daily	Uniform
Documents Acquisitioned	0	50	Daily	Normal (mean 20)
Selective Dissemination of Information	Frequency = 1/month			

Figure 4. Simulation Inputs

USAGE OF RESULTS

The results of the simulation runs were analyzed and used to provide an analytical overview of the operational processes at the information center. The results show over- and under-utilization of resources, indicate which resources are most in demand, and give minimum, maximum, and average request processing times. If other statistics are desired, they can be gathered by using the STAT construct in the process descriptions.

It is important to realize that the functions of the simulation are based only on those activities modeled and may not show the entire real-world operation. For example, although the document technician use shows at only 70.1%, this does not necessarily indicate that in reality she is under-utilized. In this model, management and supervisory functions were not modeled, and the document technician performs certain supervisory functions as well as the modeled request analysis tasks. However, the results do show relative utilization and can be very helpful in indicating areas that should be further analyzed. The model can also be valuable to management in anticipating potential problem areas. Since the data entry clerk is utilized 99% of the time, even a small increase in the volume of requests will cause a backlog at data entry. By recognizing potential problems, management can develop contingency plans and take steps to alleviate problems in advance rather than having to react to problems when they have already occurred.

As new ideas develop for improving operational efficiency, they can be modeled and simulated prior to their implementation, a valuable tool for measuring the potential impact of changes. The simulation results will indicate if the new procedure will be more efficient, reduce resource utilization, improve throughput, and have impact on other areas of the operation. This could result in cost savings since ideas could be simulated and analyzed prior to the cost of their implementation.

RESOURCE UTILIZATION

RESOURCE	MEAN BUSY PERIOD	PERCENT UTILIZATION	REQUESTS
DOCUMENT TECHNICIAN	26.0	70.1	524.0
DATA ENTRY CLERK	22.5	99.0	1270.0
KEYPUNCH	18.3	52.2	921.0
COMPUTER PRODUCTION TECH.	15.9	88.4	521.0
RJE TERMINAL	15.4	82.7	516.0
DISTRIBUTION CLERK	21.1	88.9	790.0

QUEUE STATISTICS

QUEUE	WAITING TIME		QUEUE LENGTH		
	MEAN	STD DEV	MEAN	MAX	ENTRIES
DOCUMENT TECHNICIAN	163.4	91.8	7.7	30.0	464.0
DATA ENTRY CLERK	498.8	294.1	74.1	184.0	1260.0
KEYPUNCH	5.8	4.5	0.1	1.0	38.0
COMPUTER PRODUCTION TECH.	18.8	33.8	0.6	3.0	335.0
RJE TERMINAL	27.8	32.2	0.1	1.0	13.0
DISTRIBUTION CLERK	96.8	70.6	7.9	23.0	731.0

PROCESSING TIMES

	MIN TIME	MAX TIME	MEAN TIME	STD DEV	TOTAL COMPLETED
STANDARD REQUESTS	364.1	3287.1	1819.8	83.7	4432.0
SPECIAL REQUESTS	237.6	2855.9	1549.8	228.5	283.0
SUBSCRIPTION REQUESTS	379.9	3306.0	2124.3	190.8	579.0
DISTRIBUTION	2.7	467.4	18.2	1.9	4759.0
SOURCE DATA ANALYSIS	23.2	384.7	156.8	6.7	381.0

Figure 5. Simulation Outputs

CONCLUSIONS

The insight received from the simulator to date has been used to improve the information center performance factors noted previously: responsiveness, throughput, reliability, quality of information.

Although this report reflects only a single year's effort, the feedback to date has been enthusiastic. The tool is valuable to be able to demonstrate that an information system is being managed in a cost effective manner and in addition can be used to assess resource allocation and workload distribution, identify potential problems, and study work flow paths through the information center. As larger and more real time (energy management, civil disaster, military) information systems develop, the utility of this tool looks even more attractive. With more technologies emerging, the cost-benefit ratios for a given information system can be assessed by simulating usage. This will contribute to more confident management either in the proper introduction of new technologies or the preclusion of others which, although theoretically beneficial, are not viable in practice.

REFERENCES

Wong, G., "*Computer System Simulation with GASP IV,*" Proceedings of the 1975 Winter Computer Simulation Conference, December 1975.

Wong, G., McCall, J., Seyfarth, T., "*Computer Network Simulation,*" GE TIS 75CIS07, December 1975.

McCall, J., Wong, G., Seyfarth, T., Roberts, S., and Nair, P., "Information and Data System Simulation," GE TIS 76CIS12, December 1976.

Stephen A. Rubin
Battelle, Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

SERVICE CHARGING: AN IAC EXPERIENCE

Numerous articles have been written discussing the concept of value, value added and cost-benefits that are derived from having the right information in the hands of the right person in time for decision-making. Unfortunately the majority of these conclude that while there is value inherent in information, the worth is derived from the willingness and ability of organizations and individuals to pay for the services and products they receive. Conversely, if the market will not bear the price of an information system or service it may indicate that the services offered are not cost-effective, but does this indicate lack of real value of the information?

THE MCIC EXPERIENCE

This paper is based upon the experiences of the Metals and Ceramics Information Center during the past 6 years. MCIC is one of seven technical Information Analysis Centers (IACs) chartered and sponsored by the Department of Defense. Its object is to provide authoritative and comprehensive information, data and advice on the advanced metals and ceramics in support of the systems development program of the military agencies. Prior to 1972 the services and products of the Center were offered free of charge to those individuals and organizations that were operating under a Department of Defense contract. These products consisted of a current awareness bulletin, technical reports and memoranda, and handbooks. Its service consisted of providing rapid authoritative responses to highly technical inquiries. The cost of the products and services was borne by the sponsor of the Center and so were fully subsidized. During the nearly 18 years of its "free" existence, MCIC and its predecessors published 700 reports and memoranda, and responded to over 18,000 technical inquiries. Its distribution lists for its products totaled over 4000 names representing over 1100 organizations.

In early 1971 DoD formulated a policy on charges for information services by its IAC's, and MCIC began a phased transition to service charges beginning August 1971. Initially, marketing of products and billing for inquiry services were handled by the National Technical Information Service (NTIS). The objectives of the service charge program was to allow the Center to recover part-or all-of the costs associated with the output of the Center, thus providing an income to offset rising costs and to provide for expansion of services to the technical community.

The current contractual statement of work for MCIC cites that the goal for the service charge program is an achievement of an annual rate of income equal to 75 percent of the initial contract funding (710K); of this goal 50 percent is related to direct income, and 25 percent from indirect income. Direct income is defined as:

(a) Sale of publications and services which is returned directly to the contract, and

(b) Studies funded as supplements to the MCIC contract.

Indirect income accrues from separately contracted studies for both government and industrial sponsors, costs of services provided to other sponsored programs--in-house, as well as credits for the cost of nonpaying and cancelled inquiries.

The basic objective of the service charge program was to reduce the operating cost of the IAC by creating partially self-supporting IAC's. Thus, if an information resource contained 'something of value' and was responsive to its user needs, the user would be willing to pay a fair price for the value he received. This would demonstrate which IAC's were of value and would eliminate those that were non-revenue producing.

SERVICE TO SALES: INCOME

During 1974 MCIC began its own marketing program. In addition, agreements were reached with the American Society for Metals and the Mitsubichi Research Institute, in Japan, to market the various MCIC products.

Income derived from the sale of various types of products and services for the period August 1974 through April 1976 are as follows:

Product/Service	Income		
Direct Income[a]	MCIC	NTIS	Total
Publications			
Reports, Handbooks, Films	89,500	36,000	
Technical Reviews	3,500	50,000	
Inquiry Service	--	19,500	
	93,000	105,500	198,500
Special studies[b]	802,500		802,500
Indirect Income[c]	275,500		275,500
Other Creditable Income[d]	22,500		22,500
			1,299,000

(a) Funds received by the contract from sales of documents, services, and special studies/tasks.
(b) Special studies/tasks assigned to MCIC contract.
(c) Additional activities which were conducted by MCIC and Battelle personnel under separate contracts.
(d) Creditable income associated with MCIC services which were provided without charge.

The total income of $1,299,000 represents 105 percent of the base funding of $1,235,500 for the 21 month contract period. Or in terms of contract definition 81 percent of the recovery was attributable to direct income and 24 percent to indirect. In general, the overall income for the contract Period substantially exceeds the contractual goal for the service charge program.

It is obvious that in order to fulfill its goal of cost-recovery, MCIC has had to expand its activities into the special-task areas. During the period of concern a total of 20 special studies/tasks were assigned to the MCIC contract for a total of $802,500. Six additional tasks, for a total of $275,500 were conducted by the MCIC and Battelle personnel under separate contract. This represents 84 percent of MCIC income. Activities related to output products and inquiry services were responsible for 16 percent of the income.

A comparison of MCIC's income over the past four years, is as follows:

Total Income	FY75-76	FY74	FY73
Direct	$1,000,500	$314,000	$159,000
Indirect*	298,500	280,500	206,500
	$1,299,000	$594,500	$365,500
Basic Contract Funding	$1,235,500	$710,000	$716,000
Income Recovery Percent of Funding			
Direct	81.0	44.2	22.2
Indirect	24.1	39.5	28.8
	105.1	83.7	51.0
Product Income (Direct)			
Publications	$102,500**	$ 94,000	$74,000
Inquiry Service	11,000**	12,000	7,500
	$113,500	$106,000	$81,500

* Includes Creditable Income
** Adjusted to 12 month period for comparison

The total income, both direct and indirect, was increased significantly during the 1975-1976 time frame. This was attributed to a substantial increase in direct income, as represented by the special studies requested by various government agencies.

PUBLICATIONS AND SERVICES

Income to an information analysis center can be derived from three principal sources; (1) publications, (2) inquiry services, and (3) special studies. As seen above, the special studies were the major source of revenue to MCIC. The effect of imposing a service-charge policy on an IAC, however, does not affect this activity. Since in all cases, it is paid for by additional appropriations.

Publications

This category includes state-of-the-art reports, proceedings of current technical conferences, handbooks, and current awareness publications. During the approximately four years that the MCIC sales program has been in existance, 11,929 copies of the 31 technical reports available have been sold. This has created a revenue of $150,000. Additionally, 4,400 copies

of the 14 handbooks and their supplements have been sold for a recovery of $155,000. For the technical reports, this represents a net cost recovery of 39 percent based on the preparation and publication costs. It should be noted that the investment in reports during any given fiscal year will not be recovered until future years.

The sale of the 14 handbooks since their first publication has recovered over 47 percent of the cost of preparation and publication. These have shown a much healthier return on investment than have the state-of-the-art reports.

Price is a critical factor in the marketing of an IAC product. The pricing objective of the IAC, to provide significant advanced information in a rapid, concise form, may be compromised in order to achieve marketability. The choice of topics for publication may reflect the constraint of recovery costs; it is imperative however, that an IAC continue to maintain its integrity and authoritative voice, while at the same time, expand its user community. Care must also be taken not to reduce the quality of output while doing this.

Nevertheless, the cost of preparing and publishing a report or handbook is a known quantity. It costs X dollars! A well defined and cultivated user community such as MCIC's can be "guesstimated" to purchase a known quantity of each publication during the useful life of the document. Therefore, a price can be set. It is to be remembered that the DoD service-charge policy was originally promulgated to reduce direct government funding to support IAC's to provide additional funds which could be diverted to additional output products and services. The considerations facing a commercial publisher regarding marketability and profits were not inherent in the DoD policy.

Under the free service period, each MCIC registered user received a copy of each product published in his particular field of interest. The initial distribution of such technical publications was 2500 copies. Under a service charge, averaging $13 per copy, the number of state-of-the-art reports sold ranged from approximately 50 to a little more than 1200.

We may be deluding ourselves by thinking our market is greater than these figures reflect. However, the MCIC mailing list which is revised each year, contains the names of approximately 5000 persons and organizations that have either contacted us or have been referred to us.

Inquiry Service

The philosophy of the management of MCIC has been based upon the belief that the most important service to be performed is to assist in the solution of immediate, urgent technical problems. Thus, fast and authoritative responses to urgent incoming requests for technical assistance has always been given high priority. MCIC is quite unique in this respect. Located within the Battelle-Columbus structure, users of the Center benefit significantly from the technical expertise immediately available in metals and ceramics technology, and from the several materials data banks available to the MCIC staff. An inquirer has been able to discuss his problem with, and ask questions of, a contempory technologist with the experience and resources necessary to assist in rapid solution of the technical issue.

The greatest impediment to increased technical inquiry activity appears to be the user-charge concept. The engineer/scientist in both government and industry has difficulty in negotiating payment by his organization and often the time required to complete such negotiations reduces and even negates the timeliness of the answer being sought.

Figure 1 presents a rather graphic picture of what happened when a service charge was placed on the inquiry service. In the period 1968 to 1972, MCIC had handled an average of approximately 100 technical inquiries per month. These ranged from bibliographic requests to complex highly technical problems that required multidisciplinary teamwork to provide a response. There was no limit placed on an organizations utilization of MCIC, only that it be working under a United States government program.

Since the initiation of the service charge in January 1972, and after an initial shake-down period of adjustment MCIC has received a total of 1110 requests for technical assistance during the past 45 months, or a little more than 24 inquiries per month. This alone constitutes a reduction of 76 percent in the number of monthly inquiries received.

Moreover, of the 1110 inquiries received, 598 were processed and 512 were cancelled. In general, the cancellations resulted primarily from limitations imposed on the requestor by the level of fee assigned by MCIC and/or the inability to authorize expenditures of funds. Requestors must be informed of charging requirements and methods of implementing payment at the time that a request is made. Usually, while the requestor is determining whether his company (supervisors and/or business administrators) will allow charges, MCIC must evaluate the request and estimate its cost. Consequently, most requests require two or more telephone contacts, often including one with a company procurement official, and sometimes a letter explaining the type of answer to be given and the cost. If authorization is received, usually by way of a delaying purchase order, the MCIC response must always include a written reply, often with copies of the covering letter going to the requestor's business office, his supervisor, or both. The burden of "red tape" required both by MCIC and the MCIC user tends to counteract the Center's capability of instant response to urgent technical requests, causes more than half of the incoming requests to be cancelled before authorization to proceed can be obtained, and results in costly, unreimbursed efforts by both MCIC and user.

Of the 598 assigned requests, 230 were subject to a user charge and 368 were answered free of charge. In general, the free service inquiries represented either a minimal effort by the MCIC staff, nominally less than one hour. Other 'freebies' were processed on a quid pro quo basis.

Net income to MCIC of about $85,000 has been realized from the 230 inquiries subjected to the service charge, or an average of $370 per inquiry. It is interesting to note that user charges ranged from $15 to $7500. Most government agencies and their contracting industries, particularly the aerospace industry, have no simple mechanism for paying for rapid responses to technical requests. Thus, MCIC has originated a task order

contract plan which permits an agency or company to maintain an open-ended agreement through Battelle for servicing all technical inquiry needs. Direct billing to either the organization or its NTIS deposit account are also prescribed methods of payment.

Again, it may be of interest to know who the requestor of information were, and what general classification of information they were requesting.

Requestor	Percent of Total Inquiries	Technical Scope	Percent of Total Inquiries
Industrial	74	Properties	43
Government	20	Processing	21
Academic	3.5	Applications	12
Foreign	2.5	Marketing	8
Type of Response		Fundamental Behavior	6
Technical	70	Test Methods	4
Bibliographic	30	Export Controls	3
Materials Scope		Miscellaneous	3
Metals	84		
Ceramics	13		
Composites	3		

The assignment of changes for individual specific inquiries raises questions that are quite difficult to answer. Does one charge on the basis of time spent, and if so, how is this determined? A request for coverage of a broad field might be quickly answered by an existing bibliography (already paid for by someone else), while the answer to a unique, if somewhat trivial technical question, might involve many hours to answer. Many problems, most of them predictable resulted from initiation of charges for information services. Applying a charging scale or "yardstick" to answers to inquiries "before the fact" has been particularly troublesome because a cursory effort must be expended to determine (1) whether a simple file search will suffice, or (2) whether the services of a Battelle technologist must be used, and the initial effort often may yield a correct answer, positive or negative, before authorization to proceed has been given. Is a negative response subject to a user-charge and how may users pay for negative or no information? Experience has taught us that a negative answer, regardless of the effort necessary to determine it, is not marketable. These are very real practical questions that cannot be answered by economic theories.

Originally MCIC had contact with the technical person who had direct need for the information center's service. With the advent of the service charge, the line of communication was through the corporate or technical library who had control of the purchase of documents as well as information services. These people are reluctant to have their budgets made available to outside sources. They feel that they as the reference librarian can answer any technical question. There are some cases where that is true, especially with the advent of the commercial on-line services. However, if a true technical evaluation or analysis is required, these people just cannot provide the proper response.

Now that corporate authority is required for the utilization of an IAC, utilization by the technical staff cannot be kept a secret. Management will know where the technical support has come from and will understand that this type of information service costs money. Another positive side is that the persons requesting technical assistance, since they are paying for the service, really need and will utilize what they receive. When something is free everyone wants it regardless of a requirement.

MCIC caters basically to the aerospace industry and the imposition of the service charge corresponded to a significant slowdown in defense-oriented expenditures by the Federal Government. This has particularly hurt the MCIC activity. That coupled with a prolonged period of economic recession has created a difficult and unnatural situation where organizations have been endeavoring to keep their information budgets in-house.

CONCLUSIONS AND RECOMMENDATIONS

The true IAC user community has now been identified by its willingness to pay for what is received. The audience is smaller than was expected and desired, but is still large enough to support an IAC. This has created a need for a new concept in IAC operations, the need for an agressive marketing program. Since most IAC's are technical in nature, they have been unknowledgeable and therefore unsuccessful in dealing with the promotional and marketing concepts. The arrangements with NTIS have been generally unsuccessful due to that agency's shotgun approach to promotions. The IAC product is a specialized one and must be aimed at a specific technical customer. The IAC itself has the greatest knowledge of its scope, its products, its users and their specific needs. It must now learn how to more effectively "sell" itself to the user audience.

A standard national policy for service charges pertaining to science and technology information analysis centers still does not exist. Some agencies have developed their own criteria for what they consider to be acceptable service charges and a number of IAC's have independently developed their own service charge system. It is now time that such a national guide-line be established.

It has often been said that the price placed on an information product or service should be equal to its value or worth. While in the whole, I agree with that statement, I must honestly say to those theoreticians who ascribe to that postulate, "Come into the real world, where survival is the name of the game". Ask any information broker about the pricing scheme they use. It is in no way based upon worth.

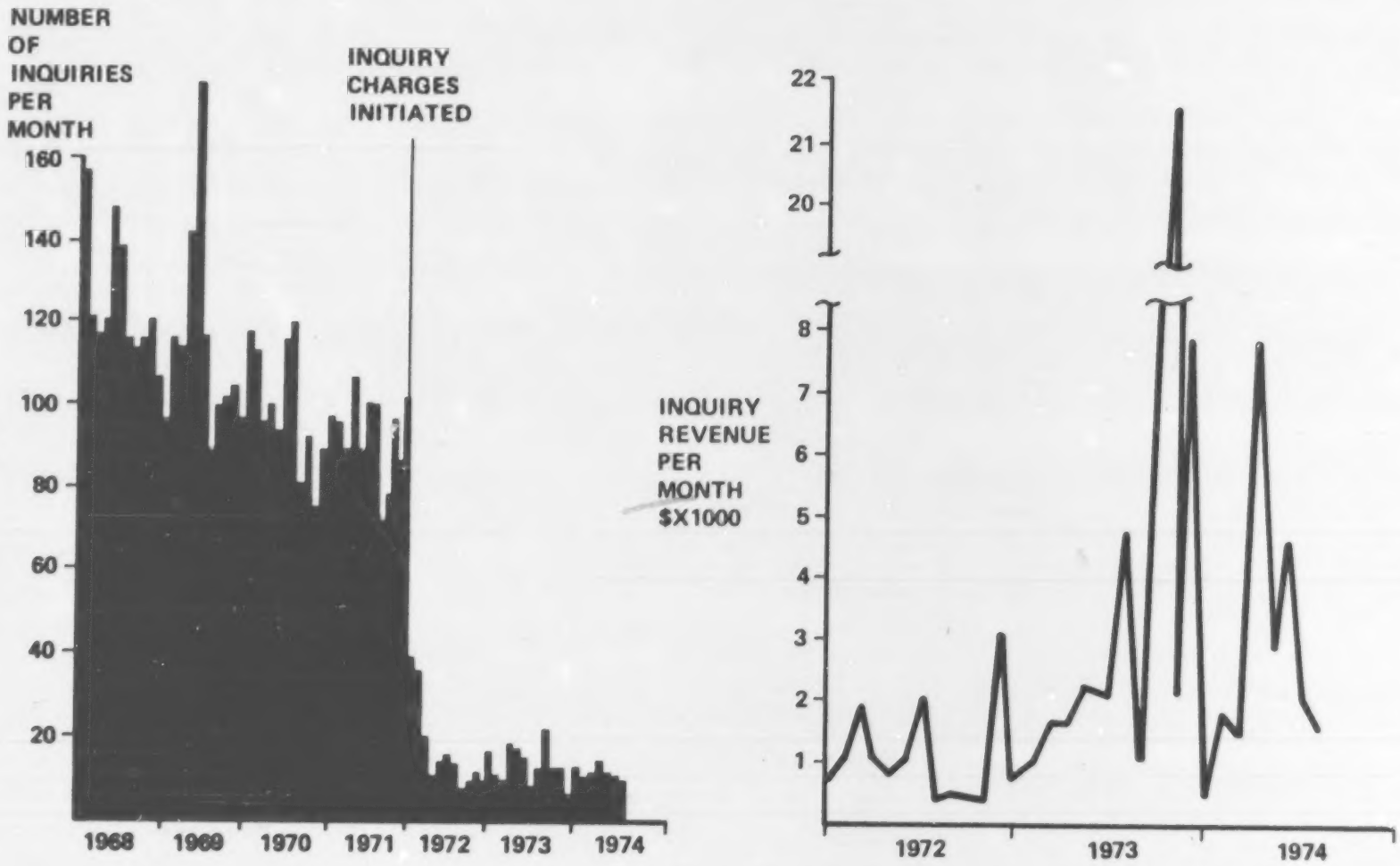


FIGURE 1. MCIC TECHNICAL INQUIRY SERVICE

185

Costs and Values of Technical Information in Industry

Philip M. Sherman
Xerox Corporation
Rochester, New York

Technical information has attributes of both cost and value, both of which are partially measurable and partially elusive. Ordinarily, we are only interested in information when its value exceeds its cost. This paper deals with the concepts of information cost and value and their relationship to one another. The concepts are dealt with at three levels: (a) a theoretical approach is described; (b) a practical approach is described, which deals with the value of information to a manager as he makes decisions; and (c) two specific document storage and retrieval systems are discussed.

Theoretical approach. Information has no intrinsic value; any value comes only through the influence the information has on physical events. Such influence is exerted through decision makers. For information to have value, it has to have three traits: (a) it must have surprise content, telling us something we did not already know; (b) the new information must cause us to take an action we otherwise would not have taken; (c) the resultant action must lead to a higher payoff than that expected if the action were not taken. An example will be used to illustrate the value of a market study in a decision-making situation, both when the study is 100% accurate and when it is only 80% accurate. The cost of the study is considered.

Practical approach. Today, we often derive information from formal information systems, usually computer-based. The procedures for developing such computer-based information systems (CBIS's) are more or less formalized; a standard sequence of phases is generally followed. However, before authorization is given for funding of CBIS development, a cost/benefit tradeoff analysis is invariably required. The costs to be considered are both tangible and intangible, as are the benefits (values) derived. The analysis procedure will be described in the paper. Intangible costs and benefits are often ignored because of the inherent difficulty in quantifying them, but they are important in many situations.

Actual situations. Two document storage and retrieval systems, in use at Xerox Corporation, are described. One is a mature system whose development costs approximate $100,000 to date, and whose operational costs approximate $450,000 yearly. The data bases in the system contain over 70,000 citations and are searched about 2,000 times a year. The costs per search can be approximated. The value of the searches is very difficult to quantify, though there are costs savings that can be identified. The second system is a younger system, with only 10,000 citations and for fewer searches each year. The same analysis is applied to this system.

Other issues. The manner of determining what information is needed is described. In brief, this derives from an understanding of what decisions have to be made. This, in turn, derives from an identification of one's objectives. A second issue is the manner by which these needs are transmitted to those charged with providing the information. Finally, a third issue is that of how all of this information should be paid for. These issues are discussed in this paper.

Information Costs and Values -- The Next 15 Years

Norman Nisenoff and Audrey Clayton
Forecasting International, Ltd.
Arlington, Virginia

Based upon several recent studies of trends affecting the dissemination of scientific and technical information, this paper explores the component costs and perceived values of STI information transfer, and various technological and external factors which are of concern in this context.

First, the "cost" aspect is examined in terms of an analysis of information delivery systems predicted to be available in 1980, 1985, and 1990, respectively. More specifically, a set of alternative system implementations which could be available in 1985 are selected for discussion based on their appropriateness to various combinations of user information requirements. The systems' capabilities and dollar costs are defined based upon extrapolation of historic values, modified by the effects of the introduction of innovative technologies whose feasibility has been demonstrated.

The paper then approaches the extremely controversial issue of weighing the costs of providing a desired information service, against the value of that service. Since no satisfactory direct measures of the true value of information presently exist, the studies here described have attempted to employ user perceptions of the relative importance of characteristics of the information delivery mechanism as surrogate measures of value. Such characteristics include accuracy, browsability, security, convenience, cost, timeliness (currency of information), and availability of hard copy. User perceptions of the value of these characteristics vary among segments of the STI user universe, with specific results quoted for the health community, industry, the library community, the academic community, and the financial community.

The characteristics are related to an expanded set of information delivery mechanisms. This analysis demonstrates the overall importance of cost, ease of access, and media type. Other factors are perceived to vary considerably in their priority ranking, depending upon the specific information product being considered.

Finally, predicted technological developments are examined in order to assess their potential impact upon the characteristics of high concern to the user, within the time frames 1975-1980, 1980-1985, 1985-1990.

Costs and Value of Scientific and Technical Information in a University Environment

Arleen N. Somerville
University of Rochester
Rochester, New York

University library services and activities are affected by their academic environment in several ways: specialized services, such as literature searches, when done at all, have been available for researchers, while students are guided in doing the work themselves; and bibliographic instruction has been aimed at students, primarily undergraduates. Some general university characteristics affect how information is evaluated by users: the availability of inexpensive student help; faculty reliance on non-university agencies for research support; the need for faculty to watch funds closely and their frequent unwillingness to spend money for information services; the reluctance of some faculty to recognize that some costs, such as staff time, are expensive; and the attraction to the university of researchers who prefer to develop their own research program.

Although almost all university libraries have experienced decreased budgets (in dollars or purchasing power), few have developed cost analysis techniques to permit effective documentation of budget requests. Few libraries have substantial amounts of relevant statistics or have conducted use and user surveys. Uncertainty exists concerning the issue of who should pay for library services - should libraries absorb the cost of such activities as computer searching, or should the cost be charged to the user.

In an effort to cope with lower budgets, many university libraries have eliminated staff positions. To compensate, libraries need to: review traditional staffing patterns; review job content to ensure that jobs are done at the most appropriate level; and establish and utilize a paraprofessional staff category to free librarians for other responsibilities. In the last five years, many individuals have recognized that a library must be effectively administered if it is to meet systematically the information needs of users. Libraries need to do more short- and long-range planning; relate the planning process to the budget; develop relevant statistics; emphasize staff development programs for staff at all levels; utilize staff at all levels more effectively; and assign decision-making responsibilities to the lowest appropriate levels.

Other ways of coping with decreased budgets are: 1) exploiting one's own collection to the fullest; 2) promoting and utilizing such inter-library cooperative efforts as acquisition, storage, and access either in person or through inter-library loan, with increased delivery time; 3) tailoring the selection of materials for the collection more towards meeting immediate needs, with a resulting loss in ability to accommodate shifts in research and teaching programs; and 4) acquiring special grants.

The library staff need to show faculty that relevant information can be provided at a cost-effective or at least reasonable rate and that the library can indeed serve as the information center on campus. The library staff needs to develop an active, even aggressive, program to advertise its services. The staff need to meet with the users in offices and labs to discuss information needs with

the users. This interaction between the user and the librarian as they seek to define the user's information needs and discuss potential ways to meet the needs is critical in establishing credibility and showing what it means to be an information center. The library staff also need to prove that they understand enough about the subject to do an effective job and that they know how to use the necessary resources. They need to supply information, not just documents, and do whatever is necessary to obtain the information. For example, librarians should work with faculty and students to obtain information, including literature searches, necessary in the preparation of grant proposals. By showing that such teamwork can save time and provide a high percentage of relevant references, the result should be cost-effective even if computer searches are charged back to the user.

Once the library has established itself as an effective information center, faculty may be more willing to go to the university administration in support of the library's overall information program. They may also be more willing to ask for information and not view it as "something I can get along without", or will not say "Don't bother, I don't really need it."

EQUIPMENT SELECTION CRITERIA FOR 1200 BAUD BIBLIOGRAPHIC SEARCHING

Alan K. Stewart
Standard Oil Co. (Indiana)
Amoco Research Center
P.O. Box 400
Naperville, IL 60540

Descriptors: On-line bibliographic searching, Telecommunications, Terminals, Full-duplex modems, Interactive information retrieval

ABSTRACT

Convenience of operation, restrictions on mode of on-line vendor access, print quality, noise level, cost, and adaptability to future needs evolved as the criteria for choosing 1200 baud search hardware at the Amoco Research Center. Investigation revealed widespread incompatibility and a total lack of standardization among equipment manufacturers, networks, and vendors. The most widely-used full-duplex modem, the VADIC VA3405, was ultimately selected because it allows for one-key carriage return, "break" capability, and line status checking. Using this modem, 1200 baud connection to SDC, Lockheed, New York Times, and BRS data bases is via TELENET. In addition, both SDC and Lockheed promise access through TYMNET, and Lockheed has a direct-dial telephone number.

A General Electric TermiNet 1200 impact printer was chosen due to its excellent print quality and acceptable noise level. The printer, without a keyboard, is driven by a Western Union VIDEO 100 CRT, which provides much greater visibility than the TermiNet with attached keyboard.

Costing approximately 2-1/2 times the monthly lease charge for a 300 baud thermal-printing terminal and acoustic coupler, the 1200 baud system offers more cost-effective on-line citation printing and browsing of inverted indexes. 1200 Baud searching can therefore involve more interaction with indexes and citation files and can be used to eliminate or reduce mailing delays from off-line printing. Costs and operating parameters continue to be affected by the ongoing introduction of new equipment and access modes.

THE PROBLEM

The number of on-line bibliographic searches performed in information centers and libraries has increased steadily over the last few years, and this growth can be expected to continue as search services continually seek, find, and train new users. At the same time, current users are becoming more sophisticated, not only in terms of data base knowledge and search system familiarity, but also in expecting a user-oriented approach to all aspects of information retrieval. The desire for higher search speeds at reasonable cost exemplifies this trend.

The transition from 110 baud (10 characters/second) to 300 baud (30 characters/second) is much more straightforward than converting to 1200 baud (120 characters/second); yet there are significant features of 1200 baud searching that recommend its adoption. This paper reports the development of criteria for choosing a terminal installation at the Amoco Research Center and suggests guidelines for those who become involved in a similar selection process. The sections of the paper correspond generally to the sequential steps followed: becoming aware of incentives, investigating background, developing criteria, and selecting equipment.

DEFINITIONS

Telecommunications terms[1] relevant to this discussion include:

BAUD. A unit of signalling speed equal to the number of discrete conditions or signal events per second.

1200 baud = 1200 bits/second = 120 characters/second (cps)

300 baud = 300 bits/second = 30 characters/second (cps)

DATA ACCESS ARRANGEMENT. Data communications equipment furnished by a common carrier, permitting attachment of privately owned data terminal and data communication equipment to the common carrier network. Abbr: DAA.

MODEM. Modulator-Demodulator. A device that modulates and demodulates signals transmitted over communications circuits. Syn: Data Set.

PROTOCOL. A formal set of conventions governing the format and relative timing of message exchange between two communication processes.

TERMINAL INSTALLATION. The totality of equipment at a user's installation including data terminal equipment, data communication equipment, and necessary support facilities.

TWO-WAY ALTERNATE OPERATION. A mode of operation of a data link in which data may be transmitted in both directions, one way at a time. Syn: Half-duplex operation.

TWO-WAY SIMULTANEOUS OPERATION. A mode of operation of a data link in which data may be transmitted simultaneously in both directions over two channels. Note: One of the channels is equipped for transmission in one direction while the other is equipped for transmission in the opposite direction. Syn: Full-duplex operation.

INCENTIVES

Four full-time and five part-time professional searches plus library support staff perform a good deal of searching at our industrial research information center, so faster searching had, and has, an obvious appeal. Data transmission four times as fast, however, does not result in a fourfold

reduction in search time; total times for entering commands, reading, and thinking are relatively independent of data transmission rate over 300 baud. I estimate that, for the same "average" search, 1200 baud transmission results in something over half the search time required for 300 baud. Summit's observation,[2] though, that "a user who has learned to search on a high-speed terminal tends to maintain and use the search habits thus developed when utilizing a low-speed terminal, and vice versa" indicates the existence of different search patterns for terminals of different speeds. Advantages of 1200 baud searching we hoped to realize from the outset were:

- Reduction of the number of searches having off-line printouts by printing most citations on-line, a departure from general practice;[3,4]
- Cost-effective on-line examination of inverted term indexes, so that the expense of excessive connect time is not a barrier to the frequent use of NEIGHBOR or EXPAND commands; and
- Increased freedom to "browse" files by quickly printing indexing of retrieved citations.*

These last two features are benefits of faster printing that facilitate more interaction between a user and a data base. The goal in upgrading to 1200 baud, then, was not necessarily to reduce costs but to offer expanded service.

BACKGROUND

Published information about modems[5-7] and terminals,[8-14] including articles describing installations particularly suited to on-line searching,[15-21] provided only general guidance in determining desirable 1200 baud equipment characteristics. Searchers at other organizations contacted seemed to have happened across the 1200 baud capability because the equipment was already in use or at least planned for another primary function. Early trials of demonstration equipment produced frustrating examples of imcompatibility between search services, terminals, and modems. A standard 1200 baud protocol does not exist, and the introduction of new full-duplex modems, no two of which are compatible, worsens an already confusing situation.

SELECTION CRITERIA EVOLUTION

Based on experience gained during early trials, modem characteristics became the focus of attention in investigating the sources of system/hardware incompatibilities. 1200 Baud full-duplex operation is quite similar to 300 baud searching in that it allows for a simple, one-key, carriage return, a "break" capability to interrupt output transmission, and continuous line status monitoring by the host computer. A trade-off in choosing full-duplex

* Even a full citation with abstract can be printed on-line for a cost similar to or slightly greater than the off-line charge.

over half-duplex operation involves restrictions on vendor access; at the present time, most services cannot be dialed via all three common access methods: direct, TELENET, and TYMNET. One notes also that full-duplex modems are hard-wired, not accoustical, so a data access arrangement and data telephone must be installed.

The legibility and reproducibility of the printed characters from different terminals varied considerably, as did noise level and input line visibility. Other parameters often specified when selecting equipment include screen size (or line length and paper size for non-CRT terminals), degree of portability, and availability of magnetic tape or diskette storage units.

EQUIPMENT SELECTION

SDC and Lockheed search service personnel were especially helpful in providing information about types of modems they and the networks were supporting. Maximum compatibility and the operating characteristics listed above dictated the choice of the VADIC VA3405 modem. The General Electric TermiNet 1200 printer, which prints fully-formed characters by impact rather than with a dot matrix pattern, was chosen due to its excellent quality and comparatively low noise level. The receive-only version was selected and a Western Union Video 100 CRT added to obtain greater visibility of characters being typed. (The TermiNet has also been used with the Datapoint 2200 minicomputer for library applications.[22])

An 80-character line length and the absence of storage were acceptable for a searching machine not designed for text editing or word processing. It was decided to obtain the system on a one-year lease in order to have maximum flexibility for growth and/or change as new equipment becomes available. 1200 Baud terminal installations generally lease for $220 to $300 per month, with purchase prices from about $5000 to $8000.

RECOMMENDATIONS

Others selecting a 1200 baud terminal installation should consider following some or all of these steps:

1. Get help--from someone in your own organization's data processing department, a local user group, friendly search service vendor, or network.

2. Start early--delivery times of two months or more are not uncommon. The direct access arrangement and telephone can be ordered first.

3. Analyze each individual installation for factors such as the need to reproduce on-line printout, sensitivity to terminal noise, display characteristics needed (if frequent demonstrations to groups of people are given, for example), and budgetary factors.

4. Try a borrowed or demonstration system identical to the one under investigation. This hands-on trial can provide experience with intricacies such as padding the end of a line with blanks to prevent dropping characters due to the fast terminal operations required at 1200 baud.

CONCLUSIONS

After a long equipment selection process, the introduction of 1200 baud capability has had a beneficial effect on our search operations. Searchers are reluctant to return to 300 baud equipment when the terminal installation is in use, preferring to wait until it is free. Also, plans to retain a 300 baud acoustical modem, for connection to the CRT and printer when 1200 baud access was unavailable, were not implemented. Not only has access been generally very good, but we feel that searches with any significant degree of printing are best done at 1200 baud.

To avail themselves of the increased service features of 1200 baud operation, other organizations should already be doing a good deal of searching, say 25 to 30 hours a month. Whereas switching to 1200 baud probably will not affect total search costs dramatically, results do include increased search interaction and more timely response to user needs.

ACKNOWLEDGMENT

W. G. Stanley for his helpful commentary on the manuscript in preparation.

REFERENCES

1. Neumann, A. J. 1974. A Guide to Networking Terminology. Washington, D.C.: National Bureau of Standards Technical Note 803; 1974, March, 22 p. (NTIS Report COM-75-10279).

2. Summit, Roger K. 1974. The Effect of Terminal Speed on Search Patterns in Interactive Information Retrieval. Palo Alto, CA: Lockheed Missles and Space Company, Inc. 1974, 3 p.

3. Wanger, Judith; Cuadra, Carlos A.; Fishburn, Mary. 1976. Impact of On-Line Retrieval Services: A Survey of Users, 1974-75. Santa Monica, CA: System Development Corporation; 1976. B12-B13.

4. Buckley, Jay S. Jr. 1975. "Planning for Effective Use of On-Line Systems." Journal of Chemical Information and Computer Sciences. 1975, August; 15 (3): 161-164.

5. Anon. 1976. "The Modem's The Thing." Telecommunications. 1976, December. 10 (12): 25-28.

6. Olson, Kirtland; O'Donnell, Karen. 1976. "Update: Modems." Digital Design. 1976, September; 6 (9): 26, 28, 30-32, 34.

7. Anon. 1976. "Modems." Digital Design. 1976, January; 6 (1): 77, 78, 80, 82.

8. Anon. 1977. "Technology Profile: Printers and Teleprinters." Mini-Micro Systems. 1977, January; 10 (1): 29, 30, 32, 40, 42, 43, 46, 46, 49, 52, 53.

9. Hughes, Jim; O'Neil, Edward. 1976. "There is No End to Terminals." Telecommunications. 1976, November: 10 (11): 53-57.

10. Backler, Jordan. 1976. "Raising CRT-Terminal IQs." Digital Design. 1976, September; 6 (9): 39, 42, 44, 48, 50, 52, 54, 56.

11. Himmelfarb, Martin; Backler, Jordan. 1976. "Printers." Digital Design. 1976, April; 6 (4): 36, 38, 40, 44, 46, 50, 52, 55, 56, 58, 60, 62, 63.

12. Granholm, Jackson W. 1975. "Interactive Hardcopy Terminals." Datamation. 1975, November; 21 (11): 51-61.

13. Brown, John A.; Workman, Robert S. 1975. How a Computer System Works. New York, NY: Arco Publishing Company, Inc.; 1975.

14. Axner, David H. 1974. "CRT Display Terminals." Modern Data. 1974, September; 7 (9): 42-49.

15. Kazlauskas, Edward John. 1977. "Selecting a Computer Terminal for the Library." Special Libraries. 1977 January; 68 (1): 24-27.

16. Bonn, Jane H.; Heer, Philipp R. 1976. "Terminal Equipment for On-Line Interactive Retrieval Using Telecommunications." Special Libraries. 1976, January; 67 (1): 30-39.

17. Anon. 1975. On-Line Services References Manual. Bethesda, MD: National Library of Medicine, Bibliographic Services Division; 1975, March. (NTIS Reports PB-253417, PB-244327, PB-240126). II-1 to II-4.

18. Marshall, Doris B. 1975. "User Criteria for Selection of Commercial On-Line Computer-Based Bibliographic Services." Special Libraries. 1975, November; 66 (11): 501-508.

19. Sternick, Barbara R. 1974. On-Line Communications Devices. Bethesda, MD: National Library of Medicine; 1974. 48 p. (ERIC Report ED-096940).

20. Lancaster, F. W.; Fayen, E. G. 1973. Information Retrieval On-Line. Los Angeles, CA: Melville Publishing Company; 1973.

21. Bork, Alfred M. 1972. "Terminals for Education." EDUCOM: Bulletin of the Interuniversity Communications Council. 1972 Winter; 7 (4): 15-20.

22. Waite, David P. 1974. "The Minicomputer: Its Role in a Nationwide Bibliographic and Information Network." In: Applications of Minicomputers to Library and Related Problems. F. Wilfrid Lancaster, ed. Urbana-Champaign, IL: University of Illinois; 1974. 136-157.

THE LIBRARY DOCTORATE - A SELF-ASSESSMENT OF DEGREE HOLDERS

Herbert S. White and Karen Hasenjager
Graduate Library School
Indiana University
Bloomington, Indiana 47401

Descriptors: Doctorate degree, library education, library research, doctorates in librarianship

Abstract: A questionnaire was mailed to 570 recipients of doctorate degrees in library science who achieved their degrees between 1930 and 1975, and who could be identified and located as active or formerly active in American librarianship. There was a 71% response. The questions and subsequent tabulations sought to determine the present and preferred areas of activity of these doctoral graduates, their evaluation of the doctorate degree as a factor in obtaining and in performing in their present jobs, and their self-assessment of their involvement in research, either as an investigator or as a supervisor of research done by others. The paper attempts to relate post-doctoral research activity to pre-doctoral research and to pre-doctoral professional experience, to determine whether or not predictive devices for the selection of doctoral candidates can be developed. Finally, through an examination of position advertising for academic teaching and administrative posts, the paper seeks to determine the perceived importance of a doctorate degree in library/information science or a subject area, versus the stated requirement for specific experience or competence, in the eyes of those institutions recruiting candidates, and to weigh the dangers of potential fragmentation or compartmentalization which can result from varying insistence on pre-requisites from job applicants.

In 1970 Ray and Patricia Carpenter postulated (1) an insufficient number of doctorates for the needs of the profession. Although their report is rich in statistical data describing the population of library doctorates as of 1968, it tends to base at least that conclusion primarily on the number of faculty positions in existence and about to come into existence through expanded graduate library programs, and assumes that such posts should be filled to a far greater extent by doctoral graduates. While employment listings surveyed during the last six months of 1976 do indeed bear out the Carpenter projections, in that completion of a doctorate, in most cases the library doctorate, has become (or at least is stated) as an absolute minimum (and often the only) requirement for library school faculty positions, this does not in itself answer a question of whether doctorates are required because they are necessary for work in library school faculty positions, or whether or not they are demanded because the demand can be met.

The purpose of our survey was to examine what holders of library doctorate degrees are doing, how they perceive the importance of their degree to the performance of their tasks, and to attempt a relationship between their academic credentials and their research scholarship. This last distinction is, we believe, important. The doctorate, particularly the Ph.D. most doctoral library programs grant, is considered a research degree. Its recipients are deemed qualified to conduct research, lead research by others, and to appreciate the importance of and evaluate the result of research as it affects the profession of librarianship. This advantage of the doctoral degree is generally conceded by even its severest critics. The benefits of a Ph.D. for success in teaching or in administration are less easy to substantiate. Even its supporters would generally agree that other criteria (such as communications skills and an interest in students) are equally if not more important for a teacher, and administrative success requires aptitudes already fully documented in the business and personnel management literature.

In fact, unless the dissertation itself deals with one of these topics, there is nothing in the Ph.D. regimen to affect the candidate's ability either to teach or administer. It is the doing, supervising, teaching and evaluating of research that the doctorate, particularly the Ph.D., appears to support. Since the Carpenter study, which identified 249 earned doctorates through 1968, there has been a substantial acceleration in the production of library terminal degrees. While the availability of Federal support funds in the late 1960's and early 1970's served to provide considerable impetus for this program, the acceleration can now be expected to continue almost on its own momentum, if for no other reason than the fact that a large and growing number of accredited library schools are now involved in or committed to the establishment of library doctorate programs, and can be expected to recruit vigorously for students and even to adapt standards where necessary to keep their programs alive. Although specific statistics vary, these investigators, through the examination of the Eyman list of doctoral dissertations through 1972(2) and a specific screening of Dissertation Abstracts, identified 280 such doctorates during the period in which the Carpenters reported 249. Similarly, Nancy Lane, in her 1975 doctoral dissertation(3), a work to which we will refer again later, located 289 library doctorates through 1969, while we were able to identify 308 during this same time period. Since degrees granted in one year are sometimes not reported in Dissertation Abstracts until several years later, the discrepancy can be accounted for in this manner. It is not significant in any case.

What is significant is the fact that the growth of library doctoral degrees has been startling, and that therefore current studies concerning the use of and need for the library doctorate are very much in order, These researchers were able to identify 662 library doctorates granted by graduate library school programs at accredited American library schools through 1975 (since only those 1975 dissertations reported through August 1976 in Dissertation Abstracts are included there may in fact be a few more). Of these degrees, better than half have been granted since 1969, and better than 1/4 since

1972. There was a slight drop in 1974, but a resurgence in 1975. Even at present levels, which can be expected to increase as neophyte programs get into full swing, we will cross the 1000 threshold by 1980. The 1930-50 cumulative total, which was doubled in 1959, doubled again in 1967, and again in 1973, will once again have doubled in 1980 or 1981.

It is not the purpose of these authors to enter into discussions concerning the so-called glut of academic doctorates in general. There is even some indication in recent literature(4) that this supersaturation may be coming to an end. This study will concern itself specifically with doctoral programs in library schools, and with the oft-discussed and oft-lamented phenomenon of research paucity in the library profession. The explanations frequently put forward, which include lack of support funds, lack of time, lack of identified research needs, and lack of research journals in which to publish, appear to these writers to be little more than after-the-fact rationalizations. They have some validity, to be sure, but they are also put forward as excuses in other disciplines, with as much or as little basis. If our profession is now producing swelling ranks or individuals trained for and presumably anxious to do research (for that is the implied commitment in their acceptance of the Ph.D.), then we should presumably be seeing the results of this growth in the very immediate future, and the profession should benefit from this body of knowledge. If this does not occur, then it may not be unreasonable to ask why it does not, and what our realistic requirements for doctoral graduates in fact are.

In her study, Lane sought to relate post-doctoral publication to pre-doctoral publication and pre-doctoral experience, for the 289 individuals identified in her 1930-69 sample. Her dissertation, which identifies direct correlations between pre-doctoral publication and pre-doctoral professional experience as a predictor for post-doctoral correlation, is a valuable study well worth careful scrutiny by doctoral admissions committees. At the same time, the Lane study simply counts publications, and makes no attempt to

differentiate those reporting research, a distinction particularly important in view of the heavy amount of non-research publication in the library field.

In an attempt to circumvent at least some of these shortcomings recognized by Dr. Lane, our survey sought information directly from recipients of library doctorates concerning their present posts and past experience, and their research and publication activity. The use of a survey questionnaire also makes it possible to correlate responses, to determine how activity in one area did or did not affect activity in another. At the same time, the subjectivity introduced in asking people to describe themselves, particularly in determining whether or not their publications involved research, and in evaluating the importance of the degree to their work, cannot be ignored. It will be addressed specifically particularly in the analysis of those questions in which it can be expected to provide a substantial bias.

As indicated above, this investigation identified 662 doctorates granted by library schools accredited by the American Library Association, between the years 1930 and 1975. Of these, 27 were deceased. Another 37 were foreign students who had returned to their native countries after receipt of their doctorate, and were excluded because they represented special cases of no direct bearing to American librarianship. Finally, 28 graduates could not be located, despite the excellent cooperation of alumni offices. Questionnaires were mailed to the remaining 570 library doctorates. Two of these individuals responded but disqualified themselves, because they had been retired for some time, and these individuals accordingly dropped from the sample. Of the remaining 568 questionnaires, responses were received from 403. The response rate of 71.0% matches exactly the level of response achieved by the Carpenter study, and provides ample evidence to the high degree of interest in and concern about this topic. Since the survey population was more than twice that of the Carpenter study, this effort represents, to the best of our knowledge, the most comprehensive survey of library doctorates. Further evidence of the concern and interest in this topic comes from

the fact that 31% of all respondents took the trouble to append specific comments, some with considerable elaboration. These ranged from cautions concerning the over-interpretation of self-evaluative data to high interest in learning results of the survey. Perhaps of greatest interest were the lengthy explanations for self-perceived paucity of research, despite the fact that, in a totally anonymous survey such as this one, no conclusions about individuals could be drawn. A large number of individuals felt obliged to explain what to them was obviously an unsatisfactory record.

Of 396 individuals responding to the question, 33.8% categorized themselves as being in the field of library administration, 3.8% in library operations, 51.3% in library education, and 11.1% in library research, as distinct from any of the above. It is assumed that this last grouping would include individuals in government posts, with commercial firms involved in research contracts, and in academic institutions but without teaching responsibilities.

Respondents indicated a remarkably high degree of contentment with their present areas of activity although, as pointed out above, the subjectivity of self-perception must be taken into account. Of doctorate library administrators, 76.1% in fact stated that they preferred administration to other areas, with education running a distant second at 10.9%. Of individuals in library operations, 53.3% preferred this activity, while 26.7% would prefer administration, an assignment they likely view as a promotional opportunity. Of the library educators who make up more than half of the survey, a remarkably high 86.7% express a preference for their present area of activity, with only 7.1% preferring administrative posts. Along the same lines, 67.6% of the individuals calling themselves researchers expressed a preference for their present line of work. While the above statistics indicate a high degree of job satisfaction or at least adjustment, they also warn of a high degree of compartmentalization and specialization, under which individuals who choose to teach without prior operational or administrative experience are not likely to acquire it at a later date, and under which even doctoral

graduates who choose another career path are not likely to teach in the future. Separately compiled data will show that this compartmentalization for librarians without doctorate degrees has become almost absolute, since it has become almost impossible for them to teach. As more doctoral graduates have become available, insistance on the doctoral degree as a prerequisite for teaching posts has become almost total. As a result, it can be postulated that, insofar as it is considered desirable that library school faculty have operational experience in the areas in which they teach, they must achieve it before they become doctoral candidates. If they enter faculty ranks upon receipt of their doctorates without prior operational or administrative experience, it does not appear, from this survey data, that they will be likely to acquire it later. Later statistics in this report will elaborate on this contention.

Survey respondents were asked to evaluate the importance of the doctorate both in achieving and in performing their present posts. Despite the fact that individuals who already hold the degree might be expected to put a greater value on that qualification in performing their duties than disinterested observers, clear and consistent distinctions emerge, which indicate a level of overqualification as between requirements for obtaining a post and the requirements for performing the duties. This tendency toward overqualification, which is recognized to be universally prevalent in academic settings, is presently under legal attack because of its impedance of the implementation of equal opportunity and affirmative action programs.

Respondents were asked to evaluate the doctorate in obtaining and in performing their present duties, by indicating whether they considered the degree essential, important, useful, or unimportant. With a weight factor of 3 for essential, 2 for important, 1 for useful, and 0 for unimportant, respondents as a whole gave an average weight of 2.40 to the doctorate in obtaining their present posts, 1.99 in performing it. As might be expected, library educators gave the doctorate the highest value rating, with 2,81 for obtaining their posts, and 2.33 for performing it. Nevertheless, an average difference

in rating of almost half a point from 196 library educators who responded to this question is not insignificant. Administrators ranked second, with a 2.06 rating for obtaining the position, 1.73 for performing it. As expected, personnel in library operations ranked the degree lowest, with a 1.20 rating for obtaining the position, and a 1.07 rating for performing it.

Perhaps most surprising is the fact that self-professed researchers ranked the doctorate at an average of 1.86 in obtaining their posts, and at 1.51 in performing it. Since this evaluation of what is fundamentally a research degree for the purpose of doing research ranks below the self-perceptions of administrators, let alone educators, it may be (although the survey did not seek to determine) that a greater proportion of these researchers work in other than academic settings, where the value placed on the doctoral degree is not as automatic. Survey respondents were asked whether they had taught, although not necessarily at the time of the survey) and if so, whether they had supervised the research of doctoral students. Of the respondents who answered this question, 136 had supervised doctoral students, 160 had not. The responses of these individuals to the question which elicited their views concerning the importance of the doctorate in obtaining and in performing their tasks were then evaluated. These response categories differ somewhat from those of library educators, since some of the individuals who classified themselves as teachers were only teaching on a part-time basis, and preferred to categorize themselves primarily as administrators or as researchers. Others had taught, but were no longer doing so. Nevertheless, while the specific numerical values differ from those reported earlier, the relationships remain closely consistent. Those who, as teachers, were supervising or had supervised doctoral research, rated the doctorate at a mean of 2.75 for obtaining their posts, and at a mean of 2.33 for performing the duties, and the same approximate 1/2 point self-perceived overqualification differential noted earlier appears here as well. Those teachers who did not supervise doctoral research rated the importance of the degree in obtaining their posts

at 2.30, in performing it at 2.07.

The Lane study mentioned earlier indicated a correlation between pre-doctoral experience and post-doctoral publication, and between pre-doctoral publication and post-doctoral publication. Individuals who had more experience before receiving their doctorates, according to Dr. Lane, published more after receiving it. Similarly, individuals with a record of publication prior to the receipt of their doctorate were more apt to publish afterwards. As Dr. Lane points out, her statistics were drawn simply from author listings in Library Literature, and made no attempt at evaluating the kind of literature being reported.

In our survey, correlations were drawn between pre-doctoral professional experience (both in and out of the library field), pre-doctoral published research, and post-doctoral published research. Even though the responses to these questions carry a recognized bias in that the evaluation of what constitutes research is left to the author (it can be assumed that he might be more charitable in his own evaluation of the research caliber of his work than an outsider would be) -- the positive impact of prior experience or of prior self-evaluated research publication on post-doctoral research disappears. If anything, the reverse appears to be true, although not significantly. For individuals receiving their doctorates after 1969, those reporting no pre-doctorate professional experience average 1.06 self-estimated publications per year after receipt of their doctorates, those with 1-3 years of prior professional experience reported 0.87 annual post-doctoral research publications, those with 4-6 years 0.65, those with 7-10 years 0.69, those with 11-15 years 0.93, and those with 16 or more 0.64. Since almost all respondents answered these series of questions, it can be seen not only that the sheer quantity of professional experience brought to the doctorate has little impact on later research activity, but also that this activity itself averages to something under one publication/year. No attempt was made in this survey to determine the kind of professional experience prior to the doctorate. If any correlations

or prediction models are to be drawn, such an investigation would appear highly desirable.

Attempts to correlate pre-degree research publication with post-degree research publication also failed to yield any meaningful trend data, although the number of annual publications since receipt of the degree for this sample grouping is consistently higher than in the previously reported data. One possible explanation for this discrepancy lies in the fact that about half of the survey respondents, while often willing to discuss pre-doctoral experience, were not willing or able to furnish data concerning their pre-degree research publications. For these individuals, therefore, no correlation with post-degree publication can be attempted. The hypothesis can be advanced that individuals unable or unwilling to report pre-doctoral publications might have been reluctant to do so because there was little or none to report. While on the face of it this might seem absurd in an anonymous survey, it has already been reported that a great many respondents felt constrained to explain what they felt to be an inadequate record, even though no one knew who they were. In any case, while the specific numbers differ, the absence of correlation already reported with regard to pre-doctorate experience also holds for pre-doctorate publication. Those individuals with no pre-doctorate research publications report post-doctorate published research at an annual rate of 1.33 articles. This rate drops to 0.73 for those with one pre-doctorate self-evaluated research article, and then fluctuates to indicate a rate of 0.81 for those with 2 publications, 1.13 for those with 3, 1.25 with 4, 0.74 with 5, 1.00 with 6, 0.76 with 7, and 0.87 with eight or more. Although we know that these reports are exaggerations because our annual literature is not that plentiful, it is clear that the patterns so apparent in Dr. Lane's study, which measured publication count, disappear entirely even when the author is asked to count his own research publications. Individuals who, by their own assessment, had not performed or at least not published the results of post-doctoral research, were asked to explain why.

Of the 235 who reported a self-perceived paucity in their own research, 61.7% attributed it to being too busy, 18.7% to the fact that research was not required of them, 7.2% to the fact that they lacked facilities (this was sometimes intended to mean laboratories or computer access, sometimes typist help or money for postage), 6.4% to being too new, and 6.0% to simply not being interested in doing any research. The explanation by 61.7% that failure to do more research was due to their being too busy matches the conclusions of Pauline Wilson in a recent article in the Journal of Education for Librarianship (5). In it, the author points to increasing pressure on library school faculty for research and publication, and suggests that, because of shifting priorities, teaching may have to be done less well, and participation in professional activities may have to be curtailed. Professor Wilson may be correct but these authors view the conclusions with at least some skepticism. Professor Wilson does point out that some faculty do no research simply because they aren't interested, and we would suggest that, to the 6.0% in our survey who indicated no interest and the 18.7% who did no research because it wasn't required could be added a considerable portion of the 68.9% who blame somebody else (too heavy a teaching load, no support, etc.). The argument can neither be attacked nor defended, because it is not possible to indicate whether faculty producing no research are already working as hard as they could be, or as hard as they ought to be. As Professor Wilson points out, much of the test will come with increased requirements on library faculty research for promotion and tenure, but only if at the same time safeguards against the dilution of teaching quality are taken. In library education, as in library operation, the experienced manager does not accept the choice of simplistic alternatives (in operations frequently stated as "Do you want quality or do you want quantity?"). The answer is that you want both, because, while they are related, they are not mutually exclusive. If the argument is going to be advanced that teachers are now too busy to do research, then that argument (which is certainly not unique to library

schools) needs to be substantiated, and cannot be simply accepted on faith. However, of far greater concern to this investigation is the almost 25% who have no particular interest in doing research. One wonders for what reason these individuals sought a research terminal degree, if they have no interest in further research. Cynics might respond that this amount of research and the degree are required to obtain the "union card" to permit the successful candidate to cross the magic portal and be then excused from further research. In fact, a sizeable number of respondent unsolicited comments made this and very similar points. However, if this attitude does indeed exist to any appreciable degree, it would point to a regrettable laxness on the part of library school administrations in the acceptance criteria employed for candidates, in the values which are stressed and implanted, and in the requirements which are met. Again, a disturbingly large number of unsolicited survey comments point to their own doctorates as tests of endurance rather than of academic ability.

There appears to be danger that, unless some remedial action is taken, the rapid growth of library doctorates will increase the present polarization between those who practice librarianship and those who teach it. This increasing polarization comes not only from the fact that library educators, as reported earlier, at least those with doctorates, have no desire to enter other areas of the profession, but also from the fact that the doctorate degree has become the sine qua non for posts in library education. It enjoys no such status in any other area of the library profession. By the same token, it is probable that operational experience will be less prevalent for our newer library educators, unless the schools themselves impose those requirements on their applicants for the doctoral program. The extent to which they will be willing and able to do this, while at the same time competing to keep their programs active as against an increasing number of schools also offering doctoral programs is open to question. The attractiveness of doctoral candidacy for experienced

librarians, particularly successful experienced librarians at an already high rate of pay, is questionable. Despite fellowships and other financial aids, the doctoral program is something of an endurance contest, and the degree is frequently the reward for several years of semi-starvation. Few already successful librarians can afford to or will want to put up with it.

The extent this polarization has already reached can be pointed out in the results of a check of advertisements for professional librarians over a six months period, from June 1976 through December 1976. Two kinds of positions were screened, those which requested candidates for library school teaching posts, and those which offered head administrative posts in academic libraries. It was felt that these were the two kinds of positions for which the doctorate in librarianship would be most appealing, and spot checks of other kinds of positions bear out this assumption. The requirement of a doctorate, or its desirability, was not mentioned to any significant extent for any library post other than the two types mentioned above. During this six months check of the Chronicle of Higher Education, Library Journal, College and Research Libraries, Wilson Library Bulletin, LJ Hotline, and variety of other publications, 46 library school faculty positions were advertised. Of these, 10 were in schools whose programs are not accredited by the American Library Association, 36 in programs which are accredited by ALA. Of these 36, 17 were in schools with active doctoral programs. Thirty-nine of the 46 position ads categorically stated the requirement of a doctorate of some kind. Twenty-three specified a library doctorate, 16 indicated that a subject doctorate would be an acceptable alternative. There was little difference in the kind of school and the kind of program, although it might be expected that a school with an active Ph.D. training program might have a greater demonstrated need for this qualification than one which, for example, trained largely school librarians in an unaccredited setting. Fifteen of the 17 positions at doctorate granting schools required a doctorate, although 4 would accept subject degrees in lieu of library

degrees. The other two ads indicated a strong preference for the library doctorate, but did not absolutely demand it.

Of the 19 positions at accredited non-doctorate library schools, 18 absolutely required the doctorate, although a surprisingly high number of 12 found a non-library doctorate an acceptable alternative. The remaining school indicated its preference though not requirement for a subject rather than library doctorate.

Of 10 positions in unaccredited schools, 6 absolutely required the library doctorate, the other 4 preferred but did not insist on a doctorate, and were equally divided in preference between the library and subject specializations.

Twenty-two of these 46 positions also indicated that other skills, such as teaching ability and professional experience, were required, but in no case were these spelled out with any specificity. For at least 39 of the 46 positions the posture was therefore clearly: First of all, present us with a doctorate - after that, we'll shop for the best secondary qualifications.

For the 25 university and 28 college head administrative positions advertised during this same period, the reverse tendency applied. Twenty-three of the university and 27 of the college head administrative posts specified exact requirements of administrative experience (in years), specific skills in management, budgeting, automation, public relations, etc. Only 5 of the 53 absolutely required a doctorate, and for four of them it was a subject and not a library doctorate. Another 21 expressed the preference for a doctorate, with 5 preferring the library field, 6 a subject field, and 11 accepting either. The remaining 27 job postings mentioned no doctorate at all. Again, the pattern is clear. Specific requirements of administrative and other skills are absolute, the doctorate is nice to have, but only as an extra. It is perhaps significant that while the trend for library education posts has been toward the doctorate, the trend in administrative posts has been away from the doctorate and toward a demonstrated ability to manage. The most dramatic evidence for this comes in a study by Kaser(6) which reports that in 1960 90% of ARL head librarians had achieved doctorates. In 1976 this had dropped to 15% of ARL library administrators.

The concern over projected polarization becomes evident. Unless candidates enter the doctoral programs with substantial professional and administrative credentials (and there are considerable financial reasons for discouraging this) they will be forced, largely by lack of alternative choice, into teaching posts. The major exception to this appears to be in the case of individuals already involved in academic administration who, through release time programs, attain the doctorate while continuing to work at their jobs. These individuals will indeed have the combination of administrative experience and the doctoral degree. However, with lessened pressure to achieve the doctorate to obtain promotion in a university library, the number of individuals willing to subject themselves to this discipline becomes questionable. In any case, it appears largely limited to people within commuting distance from their jobs to a Ph.D. granting library school.

The above may appear to be critical of doctoral programs per se in American library schools,but it is not so intended. Rather, it is suggested that if we are going to sharply increase the number of library Ph.D. and other doctoral graduates, we should know why we are doing it and what we expect to get out of it.

The article by the Carpenters postulated the need for library doctorates, particularly in library administration. The argument was never explained or defended, perhaps it was assumed that the reader would agree that this was obvious. However, it is not unreasonable to examine the premise. If we are going to have more library doctorates, what can we as a profession expect to get out of it?

Will it lead to better library administrators? It can certainly be argued that solid credentials in library research, particularly research into various aspects of library administration, will enhance the capabilities of an individual who also has solid administrative credentials, skills, and interests. However, it is extremely dubious that a library doctorate will, of itself, qualify an individual for an administrative post. In fact, there is the danger that such a graduate would, based on the false security of an earned doctorate, consider himself or herself ready for an administrative

post for which they are not remotely qualified. It is significant that advertisements for library administrative posts now recognize and to some extent newly recognize that the doctorate is a supplement to and not a replacement for other needed skills.

Will it lead to better library school teaching? The argument that an individual who has not only studied but applied research techniques is best qualified, and perhaps uniquely qualified, to teach those research techniques to others, and particularly to guide and supervise the dissertation efforts of new doctoral candidates, can certainly be strongly supported. However, the argument that such a degree presents a unique or complete qualification for library school teaching is certainly open to question. There is nothing specifically required in any doctoral program of which this research is aware to teach or sharpen communicative skills, or even the use of audio-visual techniques. Research ability and experience can be claimed to produce a more well-rounded teacher, but probably only if that ability continues to be applied and used. Although specific statistical studies appear to be lacking, thousands of student course evaluations, not only in library schools but also in other fields, have stressed the first requirement that the instructor know the field about which he or she is teaching, and that the instructor have the skills to communicate that information to the class. To some extent these requirements are best satisfied by experience, and to some extent communicative skills can be acquired. To a very great extent, however, such skills depend on personality factors which are not easy to change - on enthusiasm, on a reasonably good speaking voice, and perhaps most importantly on a genuine liking for people. We have difficulty in defining what specifically makes a "natural teacher", but we have no trouble recognizing one when we see one. Library school students seem to have no difficulty in doing it, either.

All of the above would not be particularly significant in the recent almost absolute imposition of the doctorate for library school teaching posts, except for the fact that there is the afore-mentioned evidence from the 6-months analysis of position advertisements, that the doctorate is now being demanded to the exclusion of all other qualifications. Library school administrators and search and screen committees certainly do not con-

sciously shortchange personality factors, communicative skills, and experience in their recruitment practices. In a parody of business management literature, all of us are seeking "bright, articulate, scholarly, and dedicated men and women, with at least one doctorate and preferably two, with 15 years of solid experience but below the age of 32, and with solidly established family roots but a willingness to accept a $14,000/year salary". When it comes down to it, however, library schools accept, if they must, the doctorate to the exclusion of all else for their teaching posts. There is evidence, presented previously in this study, that this polarization of operation/administration versus teaching and of experience versus education is increasing. Survey responses from recipients of library doctorates indicate a startling lack of mobility of post-doctorals into or out of the teaching profession.

The growing trend toward absolute insistence on the doctorate as a qualification for teaching (while simultaneously it becomes a decreasing requirement for administrative positions) is a cause for concern, when that insistence may deprive us of otherwise qualified faculty. This is not to suggest that experienced librarians without the doctorate are therefore more dedicated, more articulate, or more interested in working with students. These capabilities and skills do not come automatically, they must still be sought and insisted on by administrators and search and screen committees, regardless of other qualifications. However, the consistent replies of our survey respondents (who could be expected to be biased in favor of the need for a doctorate since it is their own major distinctive qualification) which indicate the disparity between the need for the doctorate as a job prerequisite and the doctorate as a job skill, indicates that there might be posts and tasks in library schools for which other capabilities are as important as or more important than the doctorate, and for which different priorities should be established.

The importance of the earned faculty doctorate in establishing and retaining status for the library school within its own academic setting is recognized, and cannot be

ignored, and is probably behind much of the recent upgrading in educational requirements. In fact, schools which have recently undergone ALA accreditation visits recognize the stress placed on both the number and diversity of earned doctorates as an indication of faculty competence. Nevertheless, it would appear that a balance must be struck between the emphasis on the terminal degree and the emphasis on other qualifications, where all are not equally available, and the school must be sufficiently aware of its own priorities, program objectives, and needs to determine what comes first. To require, for example, as one school has done, a doctorate to teach courses in special librarianship automatically eliminates by fiat most of the people qualified to deal with special librarianship from experience and these writers would suggest from their own experience, that speical librarianship cannot be realistically taught by someone who has not run a special library. Perhaps it is necessary for library schools to insist, as other professional schools have done, on different priorities for different courses. In Schools of Music, for example, there are differences in the background and requirements for teaching a course in the history of music during the baroque period, and for teaching an advanced master's class in the English Horn. To some extent library schools have attempted to offset these limitations by using adjunct and visiting faculty who do not possess other absolute requirements. However, such opportunities are usually limited to schools in or near large metropolitan centers, and in any case tend to fail to provide continuity and stability of employment, participation of the individual in the governance decisions of the school, and a sequence of priorities in which course teaching becomes something more than a spare time avocation.

The one area in which the substantial increase in library doctorates should reasonably be expected to cause a sharp upsurge is in both the amount and quality of library research. To the extent to which this has not occured, it may be because so many of the doctoral graduates are new (more than half, as indicated at the beginning of this

article, have had their degrees less than 5 years), it may be because, as Professor Wilson points out, opportunities for research are lacking, or it may be, as this article suspects, that many library doctoral graduates (not unlike the doctoral graduates in other fields) are not particularly interested in research and publication, at least not at the expense of other professional and personal things they like to do. The suggestion which has been made that the paucity of research is caused by the small number of library research publications almost certainly appears invalid. Journals normally arise in response to needs, not to create them. If solid library research is done, and this research is sufficiently important and sufficiently well written to warrant being purchased and read, journals will appear to publish them.

What appears necessary to the writers of this article, if a correlation between library advanced degrees and library research is to be established, is an increased emphasis on the importance of research in the doctoral program, not only as a one time requirement to be endured so that qualifications for a particular job can be met, but as a lifetime commitment which results from the desire to pursue a terminal degree, particularly the Ph.D. The existence of doctoral graduates (in the library field as in others) who admit to having no interest in research - and it can be assumed that the real number is larger than the 24.7% who acknowledge doing no research unless forced to - would appear to be a sharp indictment of the quality of present level doctoral programs - in their selection criteria, in communicating to students the conditions and responsibilities of what the terminal degree means and requires, in the school's treatment of research, and in the acceptance of lesser standards in the undertaking of research leading to the dissertation. There is no suggestion that historical or survey dissertations are undesirable or of lower quality, but the fact that only 21.3% of the survey respondents claimed any sort of even partially experimental approach in their dissertations, even in conjunction with other techniques, while over 30% chose historical topics, at least indicates a conservatism of philosophy of approach

and attitude which, unless watched carefully, could become excessively predominant. A further inquiry into the nature of research methodologies taught at Ph.D.-granting institutions is suggested by these findings, particularly since many of these programs are still in formative stages. Opportunities for Ph.D. students in other disciplines to learn research techniques are provided at two levels, explicitly within the curriculum as a course, and implicitly through actual work as an apprentice on a professor's research project. The lack of enthusiasm for and commitment to research by many library academicians reported in this survey may well curtail the amount of "laboratory" experience possible in a library school setting, and suggests the need for further study. To what extent are Ph.D. students given access and encouraged to participate in on-going research while they are students; in such capacities as research assistantships on individual faculty research grants, consultantships in research projects originating outside the library school but within the academic institution, or outside the university with governmental agencies, or private businesses? In the area of curriculum, to what extent do the Ph.D-granting institutions in library science provide courses, or encourage students to take courses outside the school, on research techniques other than survey research, on such topics as model building, linear programming, operations research, and the manipulation of standard statistical programs such as SPSS?

The most frequent justification stated in the survey responses for lack of research is that there is not sufficient time. If adequate instruction in research techniques is not provided by the library schools themselves, then not only Ph.D. students but also post-doctoral faculty members now pressured to produce research, will be forced to find the time to develop these skills in addition to and before fulfilling the research requirements now increasingly imposed for promotion and tenure.

REFERENCES

1. Carpenter, Ray L. and Patricia A. The Doctorate in Librarianship and an Assessment of Graduate Library Education. Journal of Education for Librarianship. 11: 3-45, Summer 1970

2. Eyman, David H., comp. Doctoral Dissertation in Library Science; Titles Accepted by Accredited Library Schools 1930-72. Xerox University Microfilms, Ann Arbor, 1973.

3. Lane, Nancy D. Characteristics Related to Productivity among Doctoral Graduates in Librarianship. Ph.D. Dissertation, U. of California, Berkeley, 1975.

4. New Spirit of Confidence Evident at Graduate Schools. The Chronicle of Higher Education 13:16, December 20, 1976.

5. Wilson, Pauline. Barriers to Research in Library Schools: A Framework for Analysis Journal of Education for Librarianship 17: 3-19, Summer 1976.

6. Kaser, David. The Effect of the Revolution of 1969/70 on University Library Administration. Paper to be published.

THE VALUE OF SEARCHING MULTIPLE DATA BASES

Daniel U. Wilde and Rex J. Snodgrass
New England Research Application Center
University of Connecticut
Storrs, Connecticut 06268

Until recently, a common practice for answering user inquiries for bibliographic information was to search the single, most relevant data base. Doing so kept search costs to a minimum but also restricted search coverage. Now, the New England Research Application Center (NERAC) offers both retrospective and SDI service that accesses all data bases appropriate to each specific question or problem. This service insures that no relevant material is missed due to users arbitrarily electing not to search peripheral, but helpful data files. Experience shows that both data base usage and user satisfaction increase when economic constraints are removed.

BACKGROUND

The New England Research Application Center (NERAC) was established eleven years ago at the University of Connecticut under cosponsorship of the National Aeronautics and Space Administration. As part of NASA's Technology Utilization Program, NERAC aids and promotes the transfer of technology by helping business and industry, colleges and universities, and local governments locate appropriate technical and business information. In other words, NERAC's mission is to help industry benefit by using previously developed technology.

NERAC's attempt at achieving this mission is unique in that it involves a total service approach using multiple data bases. User inquiries are handled by a staff of full-time technical specialists. Each staff member is not only trained in computerized information retrieval, but also has formal graduate education training and industrial experience in his field of expertise. Most are trained in multidisciplinary areas, facilitating the multiple data base approach.

In order to locate technical and business information, NERAC makes use of sources from throughout the world. The following data bases are available on NERAC's in-house computer:

Abstracted Business Information (Inform)
Agricola (CAIN)
Alloys Index (ASM)
Bioengineering Abstracts (EI)
Biological Previews (BIOSIS)
Chemical Abstracts Condensates (CA)
Computer and Controls Abstracts (Inspec)
Electrical and Electronics Abstracts (Inspec)
Engineering Index (EI)
ERDA Energy Information File
Food Science and Technology Abstracts (FSTA)
Government Reports Abstracts (NTIS)
Health Effects of Environmental Pollutants (HEEP)
Institute of Paper Chemistry (IPC)
International Aerospace Abstracts (NASA)
Journals in Education (ERIC)
Mechanical Engineering Information (ISMEC)
Metals Abstracts (ASM)
Physics Abstracts (Inspec)
Pollution Abstracts (PA)
Rubber and Plastics Research Abstracts (RAPRA)
Reports in Education (ERIC)
Scientific and Technical Aerospace Reports (NASA)
Searchable Physics Information Notices (SPIN)
Vocational and Technical Education Abstracts (ERIC)
World Aluminum Abstracts (WAA)

When a user's request cannot be satisfactorily answered with these in-house files, graduate students are available for manual searching in the large University library.

Person-to-person contact also facilitates technology transfer and frequently uncovers information not available in the literature. Therefore, NERAC makes active use of technologists at NASA and other Federal Research Centers. In order to stimulate this personal interchange of technology, NERAC's technical staff includes a government engineer with broad experience in government-industry liaison work. His job is to locate expert government

technologists willing to assist NERAC users in solving their technical problems. (2)

MULTIPLE DATA BASE SEARCHING

In the early days of computerized information retrieval, only a few data bases were available for bibliographic searching and it was possible for information specialists to become expert at searching those three or four files. Since processing costs were relatively high, a user would rarely authorize searching more than one file.

Today, there are at least 50 commonly available machine readable files. Nevertheless, the literature indicates that the average searcher still accesses only one or two data bases per question and that the utilized files tend to be those that the searcher already understands. In other words, searchers tend to return time and time again to the same files. Furthermore, accessing more than one or two files per question is still expensive and users remain reluctant to authorize multiple file searches.

In contrast, NERAC's multiple data base approach is designed to optimize the number of files that are searched for each question. This means that every user question automatically accesses all appropriate data bases, insuring that no relevant information is missed due to users arbitrarily electing not to search peripheral, but helpful data files. A question can be a retrospective search or a current awareness or both. This allows a user to find out what has been done in an area and then to keep up to date. Finally, the problem of understanding many data bases is handled by a team approach. If a question is multidisciplinary and requires many data bases, several technical specialists may pool their skills and work together designing the various retrieval strategies.

VALUE OF MULTIPLE DATA BASE SEARCHING

In the NERAC environment, multiple data base searching has proved most significant. As can be expected, product improvements are constantly being introduced. So, it is hard to separate individual cause and effect relationships with any quantitative certainty. Nevertheless, the multiple data base approach has had a most positive effect on user satisfaction. (3)

At NERAC, an attempt is made to answer all user questions, independent of subject or difficulty. Participants subscribe to NERAC's service expecting to receive answers to a high percentage of their questions. They are not buying computer searches or printouts. They expect NERAC to provide them with relevant citations to articles and reports that will help them solve their problems.

A priori, the more data bases NERAC accesses, the greater the chance for finding material that will answer a given question. Analysis reveals that as the number of in-house data bases has increased, the average number of files searched per question has also increased. Similarly user satisfaction as measured by client renewal rates has also improved dramatically.

A legitimate concern is that multiple data base searching will produce journal citation overlap. Rarely has this been a problem at NERAC except in obvious areas, such as NTIS distributing complete sections from other files, e.g. ERDA and NASA. In such cases, users quickly spot duplication. A very recent study by Bearman (4) at the National Federation of Abstracting and Indexing Services verifies that actual article overlap between data bases is surprisingly low.

Actually, various data bases tend to approach the same subject from different angles. Suppose that a user is interested in solving a food packaging problem. Certainly, Food Science and Technology Abstracts (FSTA) and Institute of Paper Chemistry Abstracts Bulletin (IPC) would be the primary files. But, a recent case disclosed that Chemical Abstracts Condensates, NTIS and even NASA produced relevant material.

Engineering Index has broad coverage but INSPEC and ISMEC should also be searched in most cases since they provide deeper coverage in specialized areas. INSPEC has the best coverage of the scientific computer literature, while only Inform covers the trade publications in the data processing field.

COST OF MULTIPLE DATA BASE SEARCHING

In NERAC's mode of operation, the cost of multiple data base searching is quite low. With a fixed cost computer, the unit cost per search drops as the number of searches increases. This remains true as long as there exists any excess computer processing capacity. NERAC takes advantage of this by running its computer seven days a week. The

only significant incremental costs for searching extra data bases are for royalties, computer paper and postage.

The additional time required by the technical specialists to construct multiple file strategies has been minimized by special software design. For example, about a dozen data bases can presently be searched with a single specially designed profile. Some of the forward-looking suppliers of data bases are now considering the requirements that must be met in order to bring the various data bases into a more uniformly searchable state. Such improvements greatly aid the cost-effectiveness of multiple file searching and obviate the necessity for ingenious, if sometimes complex, computer search programs.

The cost of searching would certainly increase if NERAC were required to purchase searches from any of the on-line vendors. In spite of on-line volume discounts, it is still cheaper for NERAC to process multiple data base searches in fast-batch mode using a dedicated in-house computer.

SUMMARY

Some of the benefits and problems associated with multiple data base searching with the NERAC system have been discussed. The benefits of increased coverage, as evidenced by client satisfaction (and therefore contract renewal) and an increase in the rate of addition of new clients, far outweigh small incremental costs and some citation overlap.

BIBLIOGRAPHY

1. Teitelbaum, Priscilla. Use of Multi Data Bases: A Case Study. Proceedings of the ASIS Annual Meeting, Volume 12, 1975, pp. 130-131.

2. Wilde, Daniel U., Markiw, Susan F., Snodgrass, Rex J., Accessing Information Through People Within the Federal Research Centers. Proceedings of the 5th ASIS Mid-Year Meeting, May, 1975, pp. 67-71.

3. Wilde, Daniel U., Documenting User Benefits in an Information Analysis Center. Proceedings of ASIS Annual Meeting, Volume 12, 1975, pp. 145-146.

4. Bearman, Toni Carbo. A Study of Coverage Overlap Among Major Science and Technology Abstracting and Indexing Services. February, 1977, 84 pp. Available from the National Federation of Abstracting and Indexing Services, 3401 Market Street, Philadelphia, PA 19104, $15.00.